ESTIMATION OF THE JUMP SIZE DENSITY IN A MIXED COMPOUND POISSON PROCESS.

F. COMTE[1], C. DUVAL[1], V. GENON-CATALOT[1], AND J. KAPPUS[2]

Abstract. Consider a mixed compound process $Y(t)=\sum_{i=1}^{N(\Lambda t)}\xi_i$ where N is a Poisson process with intensity 1, Λ a positive random variable, (ξ_i) a sequence of *i.i.d.* random variables with density f and $(N,\Lambda,(\xi_i))$ are independent. In this paper, we study nonparametric estimators of f by specific deconvolution methods. Assuming that Λ has exponential distribution with unknown expectation, we propose two types of estimators based on the observation of an *i.i.d.* sample $(Y_j(\Delta))_{1\le j\le n}$ for Δ a given time. One strategy is for fixed Δ, the other for small Δ (with large $n\Delta$). Risks bounds and adaptive procedures are provided. Then, with no assumption on the distribution of Λ, we propose a nonparametric estimator of f based on the joint observation $(N_j(\Lambda_j\Delta),Y_j(\Delta))_{1\le j\le n}$. Risks bounds are provided leading to unusual rates. The methods are implemented and compared via simulations. May 22, 2014

Keywords. Adaptive methods. Deconvolution. Mixed compound Poisson process. Nonparametric density estimation. Penalization method.
AMS Classification. 62M09–62G07

1. Introduction

Compound Poisson processes are commonly used in many fields of applications, especially in queuing and risk theory (see *e.g.* Embrechts *et al.* (1997), Grandell (1997), Mikosch (2009)). Nonparametric estimation of the jump size density in compound Poisson processes has been the subject of several recent contributions. The model can be described as follows. Consider a Poisson process $(N(t))$ with intensity 1, $(\xi_i, i\ge 1)$ a sequence of *i.i.d.* random variables with common density f independent of N and λ a positive number. Then, $(N(\lambda t), t\ge 0)$ is a Poisson process with intensity λ and $X^\lambda(t)=\sum_{i=1}^{N(\lambda t)}\xi_i$ is a compound Poisson process with jump size density f. The process X^λ has independent and stationary increments and is therefore a special case of Lévy process with Lévy density λf. Lots of references on Lévy density estimation are available (see Comte and Genon-Catalot (2009), Figueroa-Lopez (2009), Neumann and Reiss (2009), Ueltzhöfer and Klüppelberg (2011), Gugushvili (2012)). Inference is generally based on a discrete observation of one sample path with sampling interval Δ and uses the n-sample of *i.i.d.* increments $(X^\lambda(k\Delta)-X^\lambda((k-1)\Delta), k\le n)$. For the special case of compound Poisson process, van Es *et al.* (2007) build a kernel type estimator of f in the low frequency setting (Δ fixed), assuming that the intensity λ is known. As null increments provide no information on f, they only keep non zero increments for the estimation procedure. In Duval (2013) and in Comte *et al.* (2014), the same problem is considered with λ unknown and in the high frequency setting ($\Delta=\Delta_n$ tends to 0 while $n\Delta$ tends to infinity).

In this paper, we consider the case where the intensity λ is not deterministic but is random. The model is now as follows. Let Λ be a positive random variable, independent of $(N(t), t\ge 0)$

[1] Université Paris Descartes, MAP5, UMR CNRS 8145.
[2] Rostock Universität.

and of the sequence $(\xi_i, i \geq 1)$. Then,

$$Y(t) = \sum_{i=1}^{N(\Lambda t)} \xi_i \tag{1}$$

defines a mixed compound Poisson process (see Grandell 1997). Given that $\Lambda = \lambda$, the conditional distribution of $(Y(t))$ is identical to the distribution of $(X^\lambda(t))$. The mixed compound Poisson model belongs to the more general class of mixed effects models where some parameters are (unobserved) random variables. Mixed effects models are extremely popular in biostatistics and actuarial statistics. They are used in studies in which repeated measurements are taken on a series of individuals (see *e.g.* Davidian and Giltinan (1995), Pinhero and Bates (2000), Antonio and Beirlant (2007), Belomestny (2011)). The randomness of parameters allows to account for the variability existing between subjects. This is why here, we assume that, for a given time Δ, we have *i.i.d.* observations $(Y_j(\Delta), j = 1, \ldots, n)$ of $Y(\Delta)$ and our aim is to define and study nonparametric estimators of f. Note that, for deterministic λ, the n-sample of increments $(X^\lambda(k\Delta) - X^\lambda((k-1)\Delta), k \leq n)$ for one trajectory has exactly the same distribution as an *i.i.d.* sample $(X_j^\lambda(\Delta), j = 1, \ldots, n)$ for n trajectories at one instant Δ. Hence, the performances of estimating procedures based on *i.i.d.* data $(Y_j(\Delta), j = 1, \ldots, n)$ may be compared with those of procedures based on increments $(X^\lambda(k\Delta) - X^\lambda((k-1)\Delta), k \leq n)$ for one trajectory. In particular, rates of convergence depend on $n\Delta$ for large n, and small or fixed Δ (not too large).

To fix notations, let $(\Lambda_j, j \geq 1)$ be *i.i.d.* with distribution $\nu(d\lambda)$, let $(N_j(t), j \geq 1)$ be *i.i.d.* Poisson processes with intensity 1 independent of $(\Lambda_j, j \geq 1)$ and consider, for $\Delta > 0$, the n-sample $(Y_j(\Delta) = \sum_{i=1}^{N_j(\Lambda_j\Delta)} \xi_i^j, j = 1, \ldots, n)$ where $(\xi_i^j, j, i \geq 1)$ are *i.i.d.* with density f, and the sequence $(\xi_i^j, j, i \geq 1)$ is independent of $(\Lambda_j, (N_j(t)), j \geq 1)$. The paper is divided in two distinct parts, one is semi-parametric and the other purely nonparametric. In both parts, our approach relies on deconvolution and requires the assumption that:

(H0) f belongs to $\mathbb{L}^2(\mathbb{R})$.

In Section 2, we assume that the unobserved random intensities Λ_j's have an exponential distribution with parameter μ^{-1} (expectation μ). Noting that $q_\Delta := \mathbb{P}(Y(\Delta) \neq 0) = \mu\Delta(1 + \mu\Delta)^{-1}$, the unknown parameter μ is estimated (see Proposition 2.1) from

$$\hat{q}_\Delta = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}_{Y_j(\Delta) \neq 0}. \tag{2}$$

Then, we define two different nonparametric estimators of f by a deconvolution approach. First, introducing

$$Q_\Delta(u) = \mathbb{E}(e^{iuY(\Delta)} \mathbf{1}_{Y(\Delta) \neq 0}), \quad \phi_\Delta(u) = \mathbb{E}e^{iuY(\Delta)}, \tag{3}$$

we observe that the Fourier transform $f^*(u)$ of f satisfies $f^*(u) = Q_\Delta(u)/(q_\Delta \phi_\Delta(u))$, a relation which is specific to the case of Λ having an exponential distribution. We deduce an estimator $\hat{f}^*(u)$ of $f^*(u)$ based on empirical estimators of $Q_\Delta(u), \phi_\Delta(u), q_\Delta$ where we have to deal with the fact that $q_\Delta, \phi_\Delta(u)$ appear in the denominator of $f^*(u)$. Then, by Fourier inversion, we build a collection of nonparametric estimators $\hat{f}_m(x)$ of f associated with a cut off parameter m. Proposition 2.2 gives the bound of the $\mathbb{L}^2$-risk of the estimator with fixed cut off parameter. Afterwards, we propose a data driven selection $\hat{m}$ of m and prove that the corresponding estimator $\hat{f}_{\hat{m}}$ is adaptive (Theorem 2.1). The risk bounds are non asymptotic.
However, if Δ gets too small, the previous method deteriorates as q_Δ becomes small and $1/q_\Delta$

is badly estimated (this is pointed out on simulations). This is why we investigate a second method which performs well for small Δ. The idea of the second method relies on the fact that for $\mu\Delta < 1$, the following series development holds:

$$f^*(u) = \sum_{k\geq 0} (-1)^k (1+\mu\Delta)(\mu\Delta)^k (g^*_\Delta(u))^{k+1}, \tag{4}$$

where g_Δ is the conditional density of $Y(\Delta)$ given $Y(\Delta) \neq 0$. Therefore, we define an estimator of $g^*_\Delta(u)$ which leads to an estimator of $f^*(u)$ by truncating the series (4) and plugging the estimators of μ and of $g^*_\Delta(u)$. Afterwards, we proceed by deconvolution and adaptive cut off (see Proposition 2.5, Theorem 2.2). The method relies on tools developed in Chesneau *et al.* (2013) and is comparable to the one developed in Comte *et al.* (2014) for non random intensity.

In Section 3, we no longer assume that Λ has exponential distribution. We enrich the observation and assume that, in addition to $(Y_j(\Delta))$, the sample $(N_j(\Lambda_j\Delta), j = 1, \ldots, n)$ is observed. In Comte and Genon-Catalot (2013), the sample $(N_j(\Lambda_j\Delta), j = 1, \ldots, n)$ is used to estimate the density of Λ. Note that a similar problem in the Lévy case is studied in Belomestny and Schoenmakers (2014). Here, we focus on estimating f *without using any estimator for the distribution of* Λ. We do not assume that Λ admits a density and the method works for deterministic (unknown) λ. The idea is the following. Assuming that $f^*(u) \neq 0$ for all u, $\mathbb{E}(\Lambda + |\xi|) < +\infty$ and $K_\Delta(u) := \mathbb{E}(\Lambda e^{iuY(\Delta)}) \neq 0$, we check that

$$\psi(u) := \frac{(f^*(u))'}{f^*(u)} = i\frac{G_\Delta(u)}{H_\Delta(u)}, \tag{5}$$

where

$$G_\Delta(u) = \mathbb{E}\left(\frac{Y(\Delta)}{\Delta} e^{iuY(\Delta)}\right), \quad H_\Delta(u) = \mathbb{E}\left(\frac{N(\Lambda\Delta)}{\Delta} e^{iuY(\Delta)}\right). \tag{6}$$

Therefore, $\psi(u)$ can be estimated by using empirical counterparts of $G_\Delta(u), H_\Delta(u)$ dealing again with the difficulty that $H_\Delta(u)$ appears in the denominator of (5). As $f^*(0) = 1$,

$$f^*(u) = \exp\left(\int_0^u \psi(v)dv\right)$$

can be estimated replacing ψ by its estimator. This yields a first estimator $\widehat{f^*}(u)$ which may not have a modulus smaller than 1. The final estimator of $f^*(u)$ is thus defined by

$$\widetilde{f^*}(u) = \frac{\widehat{f^*}(u)}{\max(1, |\widehat{f^*}(u)|)}.$$

Afterwards, we proceed by deconvolution to define a collection of estimators $\widetilde{f}_m$ depending on a cut off parameter m. Proposition 3.1 gives the bound of the $\mathbb{L}^2$-risk of $\widetilde{f}_m$ for fixed m. The risk bounds are non standard as well as the proof to obtain them and give rise to unusual rates on standard examples when making the optimal bias-variance trade-off. We propose an heuristic penalization criterion to define a data-driven choice of the cut off parameter. A naive procedure for estimating f is also described.

Section 4 illustrates our methods on simulated data for different examples of jump densities f and of distributions for Λ. It appears clearly that the first method of Section 2 performs well for all values of Δ except very small contrary to the second one, as expected from theoretical results. The method of Section 3 though more complex can be easily implemented. A complete discussion on numerical results is given. Section 5 contains proofs. In the Appendix, auxiliary results needed in proofs are recalled.

2. Semi-parametric strategy of estimation

In this Section, we assume that Λ has an exponential distribution with parameter μ^{-1}.

2.1. **Estimation of μ when Λ is $\mathcal{E}(\mu^{-1})$.** The following assumption is required:

(H1) The parameter μ belongs to a compact interval $[\mu_0, \mu_1]$ with $\mu_0 > 0$.

For any distribution $\nu(d\lambda)$ of Λ_j, the distribution of $N_j(\Lambda_j\Delta)$ is given by:

$$\mathbb{P}(N_j(\Lambda_j\Delta) = m) = \int_0^{+\infty} e^{-\lambda\Delta}\frac{(\lambda\Delta)^m}{m!}\nu(d\lambda), m \geq 0$$

When Λ_j has an exponential density $\mu^{-1}e^{-\lambda\mu^{-1}}1_{\lambda>0}$, the computation is explicit. For $m \geq 0$,

$$\mathbb{P}(N_j(\Lambda_j\Delta) = m) \quad = \quad \left(\frac{\mu\Delta}{\mu\Delta+1}\right)^m \frac{1}{\mu\Delta+1} := \alpha_m(\mu, \Delta). \tag{7}$$

Noting that

$$\mathbb{P}(Y_j(\Delta) \neq 0) = \mathbb{P}(N_j(\Lambda_j\Delta) \neq 0) = 1 - \alpha_0(\mu, \Delta) = 1 - \frac{1}{1+\mu\Delta} = \frac{\mu\Delta}{1+\mu\Delta} := q_\Delta, \tag{8}$$

we get

$$\mu = \frac{1}{\Delta}\frac{q_\Delta}{1-q_\Delta}.$$

Therefore we consider the empirical estimator $\hat{q}_\Delta$ of q_Δ given by (2). Assumption **(H1)** implies that $q_\Delta \in \left[q_{0,\Delta}, q_{1,\Delta}\right]$ with

$$q_{0,\Delta} = \frac{\mu_0\Delta}{1+\mu_0\Delta}, \quad q_{1,\Delta} = \frac{\mu_1\Delta}{1+\mu_1\Delta}.$$

To estimate $1/q_\Delta$ and μ, we define

$$\frac{1}{\tilde{q}_\Delta} = \frac{1}{\hat{q}_\Delta}\mathbf{1}_{\hat{q}_\Delta \geq q_{0,\Delta}/2}, \quad \tilde{\mu} = \frac{1}{\Delta}\frac{\hat{q}_\Delta}{1-\hat{q}_\Delta}\mathbf{1}_{1-\hat{q}_\Delta \geq (1-q_{1,\Delta})/2}. \tag{9}$$

These definitions require the knowledge of μ_0, μ_1, however this can be avoided, see Remark 1 and Remark 3. The following properties hold.

Proposition 2.1. *Under* **(H1)**, *and $n\Delta \geq 1$, the estimators $\hat{q}_\Delta$, $1/\tilde{q}_\Delta$ and $\tilde{\mu}$ given by (2) and (9) satisfy for all integer $p \geq 1$,*

$$\mathbb{E}(\hat{q}_\Delta - q_\Delta)^{2p} \leq C(p, \mu_1)\left(\frac{\Delta}{n}\right)^p, \quad \mathbb{E}\left(\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta}\right)^{2p} \leq C'(p, \Delta)\left(\frac{1}{n\Delta^3}\right)^p,$$

and

$$\mathbb{E}(\tilde{\mu} - \mu)^{2p} \leq \frac{C''(p, \Delta)}{(n\Delta)^p} \tag{10}$$

where $C(p, \mu_1)$ only depends on p, μ_1, $C'(p, \Delta) = C'(p) + O(\Delta)$, $C''(p, \Delta) = C''(p) + O(\Delta)$, and $C'(p, \Delta)$, $C''(p, \Delta)$ only depend on p, μ_0, μ_1, Δ.

We can see that for fixed Δ, the estimation rate for μ is of order $n^{-1/2}$, the standard parametric rate.

2.2. **Notation.** The following notations are used below. For $u : \mathbb{R} \to \mathbb{C}$ integrable, we denote its $\mathbb{L}^1$ norm and its Fourier transform respectively by

$$\|u\|_1 = \int_{\mathbb{R}} |u(x)|dx, \quad u^*(y) = \int_{\mathbb{R}} e^{iyx} u(x)dx, \quad y \in \mathbb{R}.$$

When u, v are square integrable, we denote the $\mathbb{L}^2$ norm and the $\mathbb{L}^2$ scalar product by

$$\|u\| = \left(\int_{\mathbb{R}} |u(x)|^2 dx \right)^{1/2}, \quad \langle u, v \rangle = \int_{\mathbb{R}} u(x)\overline{v}(x)dx \quad \text{with } z\overline{z} = |z|^2.$$

We recall that, for any integrable and square-integrable functions u, u_1, u_2, the following relations hold:

$$(u^*)^*(x) = 2\pi u(-x) \text{ and } \langle u_1, u_2 \rangle = (2\pi)^{-1} \langle u_1^*, u_2^* \rangle. \tag{11}$$

The convolution product of u, v is denoted by: $u \star v(x) = \int_{\mathbb{R}} u(y)\bar{v}(x-y)dy$.

2.3. **Estimation of f for fixed sampling interval Δ.** In this section, we propose an estimator of f assuming that Δ is fixed (heuristically, $\Delta = 1$). The distribution of $Y_j(\Delta)$ is given by:

$$\mathbb{P}_{Y_j(\Delta)}(dx) = \alpha_0(\mu, \Delta)\delta_0(dx) + \sum_{m \geq 1} \alpha_m(\mu, \Delta) f^{\star\, m}(x)dx,$$

where $\alpha_m(\mu, \Delta), m \geq 0$ is the distribution of $N_j(\Lambda_j \Delta)$ when Λ_j has exponential distribution with expectation μ, *i.e.* the geometric distribution with parameter $1/(1 + \mu\Delta)$. AS null values bring no information on the density f, we intend to base our estimation on non null data. Let us note that Q_Δ, ϕ_Δ defined in (3) satisfy

$$\phi_\Delta(u) \quad = \quad \int_0^{+\infty} \mu^{-1} e^{-\lambda/\mu} \exp\left(-\lambda\Delta(1 - f^*(u))\right) d\lambda = \frac{1}{1 + \mu\Delta(1 - f^*(u))}.$$

Simple computations yield:

$$Q_\Delta(u) = \phi_\Delta(u) - \frac{1}{1 + \mu\Delta} = \frac{\mu\Delta f^*(u)}{(1 + \mu\Delta)(1 + \mu\Delta(1 - f^*(u)))}.$$

Solving for $f^*(u)$ yields the following formula

$$f^*(u) = \frac{1 + \mu\Delta}{\mu\Delta} \frac{Q_\Delta(u)}{Q_\Delta(u) + \frac{1}{1+\mu\Delta}} = \frac{Q_\Delta(u)}{q_\Delta \phi_\Delta(u)}.$$

This formula which is very specific to the case of Λ_j having exponential distribution with expectation μ suggests to estimate $f^*(u)$ as follows:

$$\hat{f}^*(u) = \frac{\hat{Q}_\Delta(u)}{\tilde{q}_\Delta \widetilde{\phi}_\Delta(u)} \tag{12}$$

where $1/\tilde{q}_\Delta$ is defined by (9),

$$\hat{Q}_\Delta(u) = \frac{1}{n} \sum_{j=1}^{n} e^{iuY_j(\Delta)} \mathbb{1}_{Y_j(\Delta) \neq 0}, \tag{13}$$

and for k a constant ($k = 0.5$ in the simulations),

$$\frac{1}{\widetilde{\phi_\Delta}(u)} = \frac{1}{\widehat{\phi_\Delta}(u)} \mathbb{1}_{|\widehat{\phi_\Delta}(u)| \geq k/\sqrt{n}}, \quad \widehat{\phi_\Delta}(u) = \frac{1}{n} \sum_{j=1}^{n} e^{iuY_j(\Delta)}. \tag{14}$$

Then, we apply Fourier inversion to (12), but as $\hat{f}^*$ is not integrable, a cut off is required. We propose thus

$$\hat{f}_m(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\hat{f}^*(u)du. \tag{15}$$

Then we can bound the mean-square risk of the estimator as follows.

Proposition 2.2. *Assume that Λ is $\mathcal{E}(\mu^{-1})$ and that* **(H0)**-**(H1)** *hold. Then the estimator $\hat{f}_m$ for $m \leq n\Delta$ given by (15) and (12) satisfies*

$$\mathbb{E}(\|\hat{f}_m - f\|^2) \leq \|f - f_m\|^2 + \frac{1}{\pi n q_\Delta}\int_{-\pi m}^{\pi m} \frac{du}{|\phi_\Delta(u)|^2} + \frac{c}{n\Delta}$$

where f_m is such that $f_m^* = f^*\mathbf{1}_{[-\pi m, \pi, m]}$ and where the constant c does not depend on n nor Δ.

We can see that the bias term $\|f - f_m\|^2$ is decreasing with m while the variance term, of order m/n, is increasing with m; this illustrates that a standard bias-variance compromise has to be performed. We can also see this in an asymptotic way if we assume that f belongs to the Sobolev ball defined by

$$\mathcal{S}(\alpha, L) = \{f \in \mathbb{L}^2(\mathbb{R}), \int |f^*(u)|^2(1+u^2)^\alpha du \leq L\}.$$

Then

$$\|f - f_m\|^2 = \frac{1}{2\pi}\int_{|u|\geq \pi m} |f^*(u)|^2 du \leq \frac{L}{2\pi}(1+(\pi m)^2)^{-\alpha} \leq c_L m^{-2\alpha}.$$

Therefore, we find that, for $m = m^* \asymp (n\Delta)^{1/(2\alpha+1)}$, $\mathbb{E}(\|\hat{f}_{m^*} - f\|^2) = O((n\Delta)^{-2\alpha/(2\alpha+1)})$, which is a standard nonparametric rate.

We propose a data driven way of selecting m, and we proceed classically by mimicking the bias-variance compromise. Setting $\mathcal{M}_n = \{1, \dots, n\Delta\}$, we select

$$\hat{m} = \arg\min_{m\in\mathcal{M}_n}\left(-\|\hat{f}_m\|^2 + \widehat{\text{pen}}(m)\right) \quad \text{where } \widehat{\text{pen}}(m) = \kappa\frac{1}{\tilde{q}_\Delta}\frac{1}{2\pi n}\int_{-\pi m}^{\pi m}\frac{du}{|\widetilde{\phi_\Delta}(u)|^2}.$$

Then we can prove the following result

Theorem 2.1. *Assume that Λ is $\mathcal{E}(\mu^{-1})$ and that assumption* **(H0)**-**(H1)** *hold. Then for $\kappa \geq \kappa_0 = 96$, we have*

$$\mathbb{E}(\|\hat{f}_{\hat{m}} - f\|^2) \leq c \inf_{m\in\mathcal{M}_n}\left(\|f - f_m\|^2 + \frac{\kappa}{2\pi n q_\Delta}\int_{-\pi m}^{\pi m}\frac{du}{|\phi_\Delta(u)|^2}\right) + \frac{c'}{n\Delta}$$

where c is a numerical constant ($c = 4$ would suit) and c' depends on μ_0, μ_1 and $\|f\|$.

The bounds of Proposition 2.2 and Theorem 2.1 are nonasymptotic and hold for all n and Δ. However, if Δ gets too small, the method deteriorates because q_Δ is small and $1/q_\Delta$ is badly estimated.

Remark 1. Note that the knowledge of μ_0 is required for $1/\widetilde{q}_\Delta$ but we may get rid of this condition by defining $1/\bar{q}_\Delta = (1/\hat{q}_\Delta)\mathbf{1}_{\hat{q}_\Delta \geq k\sqrt{\Delta/n}}$ for some constant k. Following the proof of Lemma 5.1, we would get

$$\mathbb{E}(|\frac{1}{\bar{q}_\Delta} - \frac{1}{q_\Delta}|^{2p}) \leq c\left(\frac{1}{q_\Delta^{2p}} \wedge \frac{(n/\Delta)^{-p}}{q_\Delta^{4p}}\right) = O(\frac{1}{(n\Delta^3)^p}).$$

The results of Proposition 2.2 and Theorem 2.1 hold. This estimator, with $k = 0.5$, is the one used in simulations to avoid fixing a value for μ_0 and μ_1.

Remark 2. Theorem 2.1 states that the estimator $\hat{f}_{\hat{m}}$ is adaptive as the bias-variance compromise is automatically realized. It also states that there is a minimal value κ_0 such that for all $\kappa \geq \kappa_0$, the adaptive risk bound holds. From our proof, we find $\kappa_0 = 96$, which is not optimal. Indeed, in simple models, a minimal value for κ_0 may be computed. For instance, Birgé and Massart (2007) prove that for Gaussian regression or white noise models, the method works for $\kappa_0 = 1 + \eta$, $\eta > 0$, and explodes for $\kappa_0 = 1 - \eta$. To obtain the minimal value in another context is not obvious. This is why it is customary when using a penalized method, to calibrate the value κ in the penalty by preliminary simulations.

2.4. **Estimation of f for small sampling interval.** Now, we assume that $\Delta = \Delta_n$ tends to 0 and that $n\Delta$ tends to infinity. We use an approach for small sampling interval which is different from the previous one. The conditional distribution of $Y(\Delta)$ given $Y(\Delta) \neq 0$ has density and Fourier transform given by (see (3), (8))

$$g_\Delta(x) = \frac{1}{q_\Delta} \sum_{k\geq 1} \alpha_k(\mu, \Delta) f^{\star\, k}(x), \quad g^*_\Delta(u) = \frac{Q_\Delta(u)}{q_\Delta}.$$

Using (7), the Fourier transform of g_Δ is given by $(\mu\Delta |f^*(u)|/(1+\mu\Delta) < 1$):

$$g^*_\Delta(u) = \left(\frac{\mu\Delta}{1+\mu\Delta}\right)^{-1} \sum_{k\geq 1} \frac{1}{1+\mu\Delta}\left(\frac{\mu\Delta}{1+\mu\Delta}\right)^k (f^*(u))^k = \frac{f^*(u)}{1+\mu\Delta(1-f^*(u)))}$$

Thus $|g^*_\Delta(u)| \leq |f^*(u)|$ which implies that

$$\|g_\Delta\| \leq \|f\|. \tag{16}$$

Solving for $f^*(u)$ yields:

$$f^*(u) = (1+\mu\Delta)\frac{g^*_\Delta(u)}{1+\mu\Delta g^*_\Delta(u)}.$$

Now, if we assume that $\mu\Delta < 1$, then the development in series (4) holds. To estimate $f^*(u)$, we have to estimate $g^*_\Delta(u)$. For this, we set:

$$\widetilde{g^*_\Delta}(u) = \frac{\widehat{g^*_\Delta}(u)}{\max\left(1, |\widehat{g^*_\Delta}(u)|\right)} \text{ with } \widehat{g^*_\Delta}(u) = \frac{\hat{Q}_\Delta(u)}{\tilde{q}_\Delta}$$

where $\hat{Q}_\Delta(u)$ is given by (13) and $\tilde{q}_\Delta$ by (9). We can prove that $\widetilde{g^*_\Delta}$ satisfies the following property.

Proposition 2.3. *Under* **(H0)**, *for all $v \geq 1$ and $\mu \in [\mu_0, \mu_1]$ with $\mu_1\Delta < 1$, we have*

$$\sup_{u\in\mathbb{R}} \mathbb{E}\left(|\widetilde{g^*_\Delta}(u) - g^*_\Delta(u)|^{2v}\right) \leq \frac{C(v,\mu_0,\mu_1)}{(n\Delta)^v}.$$

We can deduce from Proposition 2.3 the following Corollary showing also that we have estimators of convolutions of g_Δ with parametric rate as soon as $k \geq 2$:

Corollary 2.1. *Let $\mu \in [\mu_0, \mu_1]$ and*

$$\widehat{g^{\star k}_{m,\Delta}}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} (\widetilde{g^*_\Delta}(u))^k e^{-iux} du, \quad g^{\star k}_{m,\Delta}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} (g^*_\Delta(u))^k e^{-iux} du.$$

Then, for all $k \geq 2$

$$\mathbb{E}\left(\|\widehat{g^{\star k}_{m,\Delta}} - g^{\star k}_{m,\Delta}\|^2\right) \leq c(k,\mu_0,\mu_1)\left(\frac{m}{(n\Delta)^k} + \frac{D_k^2}{n\Delta}\|f\|^2\right),$$

where $D_k = (3^k - 2^k - 1)/2$.

Moreover, the coefficients $F_k(\mu\Delta) := (1+\mu\Delta)(\mu\Delta)^k$ are estimated by $F_k(\tilde{\mu}\Delta)$ with $\tilde{\mu}$ defined in (9), and satisfy:

Proposition 2.4. *For all $k \geq 0$ and $\mu \in [\mu_0, \mu_1]$,*

$$\mathbb{E}(F_k(\tilde{\mu}\Delta) - F_k(\mu\Delta))^2 \leq C(k, \mu_1, \Delta)\frac{\Delta^{2k}}{n\Delta}$$

where the constant $C(k, \mu_1, \Delta) = C(k, \mu_1) + O(\Delta)$.

Thus, to estimate f^*, we plug in the estimators $\widetilde{g_\Delta^*}$ and $\tilde{\mu}$ in (4) and truncate the series up to order K:

$$\hat{f}_K^*(u) = \sum_{k=0}^{K}(-1)^k(1+\tilde{\mu}\Delta)(\tilde{\mu}\Delta)^k(\widetilde{g_\Delta^*}(u))^{k+1}.$$

Then, we proceed with Fourier inversion to obtain an estimator of f, but we have again to insert a cut off in the integral

$$\hat{f}_{m,K}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\hat{f}_K^*(u)e^{-iux}du.$$

Then, we can prove the following result.

Proposition 2.5. *Assume that* **(H0)**-**(H1)** *hold, that Λ is $\mathcal{E}(\mu^{-1})$, and that $2\mu_1\Delta < 1$ and $\Delta < 1$. Then, for any $m \leq n\Delta$, we have*

$$\mathbb{E}\left(\|\hat{f}_{m,K} - f\|^2\right) \leq \|f_m - f\|^2 + 4A(\mu_1\Delta)^{2K+2} + 12\frac{(1+\mu\Delta)^3}{\mu}\frac{m}{n\Delta} + \frac{E_K}{n\Delta},$$

where f_m is such that $f_m^* = f^*\mathbf{1}_{[-\pi m, \pi, m]}$, $A = 4\|f\|^2(1+\mu_1\Delta)^2/(1-\mu_1\Delta)^2$ and E_K is a constant depending on K, μ_0, μ_1 and $\|f\|$.

If $f \in \mathcal{S}(\alpha, L)$, then $\|f - f_m\|^2 \leq c_L m^{-2\alpha}$ as already noticed and choosing $m = m^* \asymp (n\Delta)^{1/(2\alpha+1)}$ implies

$$\mathbb{E}(\|\hat{f}_{m^*,K} - f\|^2) \leq c_1(n\Delta)^{-2\alpha/(2\alpha+1)} + c_2\Delta^{2K+2}.$$

To choose K in practice, we impose $\Delta^{2K+2} \leq 1/(n\Delta)$, *i.e.*

$$K \geq K_0 := \frac{1}{2}\left(\frac{\log(n)}{|\log(\Delta)|} - 3\right), \tag{17}$$

even if this contradicts the fact that K is fixed (and thus independent on n).

Now, we have to select m in a data driven way. To that aim, we propose

$$\hat{m}_K = \arg\min_{m\in\{1,\ldots,[n\Delta]\}}\left(-\|\hat{f}_{m,K}\|^2 + \widetilde{\text{pen}}(m)\right), \quad \widetilde{\text{pen}}(m) = \kappa'\frac{(1+\tilde{\mu}\Delta)^2}{\tilde{q}_\Delta}\frac{m}{n}.$$

We can prove

Theorem 2.2. *Assume that* **(H0)**-**(H1)** *hold, that Λ is $\mathcal{E}(\mu^{-1})$, and that $2\mu_1\Delta < 1$. Then there exists a numerical constant κ_0' such that, for all $\kappa' \geq \kappa_0'$, we have*

$$\mathbb{E}\left(\|\hat{f}_{\hat{m}_K,K} - f\|^2\right) \leq c\inf_{m\in\{1,\ldots,[n\Delta]\}}\left(\|f_m - f\|^2 + \kappa'\frac{(1+\mu\Delta)^3}{\mu}\frac{m}{n\Delta}\right) + 4A(\mu_1\Delta)^{2K+2} + \frac{E_K'}{n\Delta},$$

where c is a numerical constant, A is defined in Proposition 2.5 and E_K' is a constant depending on K, μ_0, μ_1 and $\|f\|$.

For the choice of κ' in the penalty, we refer to Remark 2.

Remark 3. A proposal analogous to the one in Remark 1 can be done here to define an estimator of μ which does not depend on μ_1: $\bar{\mu} := \hat{q}_\Delta/(\Delta(1-\hat{q}_\Delta))\mathbf{1}_{1-\hat{q}_\Delta \geq k\sqrt{\Delta/n}}$. This is used in simulations with $k = 0.5$.

3. Nonparametric strategy

In this section, we no longer assume that Λ follows an exponential distribution and turn to the estimation of f, using both samples $(Y_j(\Delta), N_j(\Lambda_j\Delta))_j$.

3.1. Definition of the estimator.

We start from the characteristic function and for ν denoting the distribution of Λ, we have, from (3),

$$\phi_\Delta(u) = \int_0^{+\infty} \exp\left(-\lambda\Delta(1 - f^*(u))\right)\nu(d\lambda)$$

and by derivation (see (6)),

$$iG_\Delta(u) = (f^*(u))' K_\Delta(u), \quad \text{where} \quad K_\Delta(u) = \mathbb{E}(\Lambda e^{iuY(\Delta)}). \tag{18}$$

For fixed $\Lambda = \lambda$, we get

$$\mathbb{E}\left(\frac{N(\lambda\Delta)}{\Delta} e^{iu\sum_{k=1}^{N(\lambda\Delta)}\xi_k}\right) = f^*(u)\mathbb{E}\left(\lambda e^{iu\sum_{k=1}^{N(\lambda\Delta)}\xi_k}\right)$$

and therefore

$$H_\Delta(u) = \mathbb{E}\left(\frac{N(\Lambda\Delta)}{\Delta} e^{iuY(\Delta)}\right) = f^*(u)K_\Delta(u).$$

We deduce that, if $H_\Delta \neq 0$, (5) holds. With the condition $f^*(0) = 1$, we obtain the formula

$$f^*(u) = \exp\left(\int_0^u \psi(v)dv\right).$$

Note that for $u \leq 0$, $\int_0^u = -\int_u^0$, and the formula is still valid. We deduce an estimator by setting

$$\widetilde{f}_m(u) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\widetilde{f^*}(u)du \tag{19}$$

where

$$\widetilde{f^*}(u) = \widehat{f^*}(u)\mathbf{1}_{\{|\widehat{f^*}(u)|\leq 1\}} + \frac{\widehat{f^*}(u)}{|\widehat{f^*}(u)|}\mathbf{1}_{\{|\widehat{f^*}(u)|>1\}} = \frac{\widehat{f^*}(u)}{\max(1, |\widehat{f^*}(u)|)},$$

with

$$\widehat{f^*}(u) = \exp\left(\int_0^u \widetilde{\psi}(v)dv\right), \quad \widetilde{\psi}(v) = i\frac{\hat{G}_\Delta(v)}{\tilde{H}_\Delta(v)},$$

$$\hat{G}_\Delta(v) = \frac{1}{n\Delta}\sum_{j=1}^n Y_j(\Delta)e^{ivY_j(\Delta)},$$

$$\hat{H}_\Delta(v) = \frac{1}{n\Delta}\sum_{j=1}^n N_j(\Lambda_j\Delta)e^{ivY_j(\Delta)}, \quad \frac{1}{\tilde{H}_\Delta(v)} = \frac{1}{\hat{H}_\Delta(v)}\mathbf{1}_{\{|\hat{H}_\Delta(v)|\geq k(n\Delta)^{-1/2}\}},$$

for some constant k.

We introduce the following assumption.

[**B**] (i) $\forall u \in \mathbb{R},\ f^*(u) \neq 0$, and there exists $K_0 > 0$ such that $\forall u \in \mathbb{R}$, $|K_\Delta(u)| \geq K_0$.

(ii)(p) $\mathbb{E}(\xi^{2p}) < +\infty$, $\mathbb{E}(\Lambda^{2p}) < +\infty$.
(iii) $\|G'_\Delta\|_1 < +\infty$.

To justify assumption **[B]**(i), let us consider the case where Λ follows an exponential $\mathcal{E}(1/\mu)$ distribution. Then

$$K_\Delta(u) = \frac{\mu}{[1+\mu\Delta(1-f^*(u))]^2} \quad \text{and} \quad K_\Delta(u) \sim_{u\to+\infty} \frac{\mu}{[1+\mu\Delta]^2}.$$

Thus H_Δ is not lower bounded near infinity contrary to $K_\Delta(u)$.

Under **[B]** (ii), $\mathbb{E}[(Y(\Delta))^{2p}] = \Delta\mathbb{E}(\Lambda)\mathbb{E}(\xi^{2p}) + o(\Delta)$. Indeed, we first compute the cumulants of the conditional distribution of $Y(\Delta)$ given Λ. Then we deduce the conditional moments using the link between moments and cumulants. Integrating with respect to Λ gives the result. Note that **[B]**(ii) implies that G'_Δ exists, with

$$iG'_\Delta(u) = (f^*)''(u)K_\Delta(u) + i(f^*)'(u)\mathbb{E}(\Lambda Y(\Delta)e^{iuY(\Delta)}).$$

Thus, if $(f^*)'$ and $(f^*)''$ are integrable, then **[B]**(iii) holds.

Then we can prove the following result.

Proposition 3.1. *Assume that* **(H0)**, [B] (i) *and* (ii)(p) *hold. Let* $\widetilde{f}_m$ *be given by (19) and let* $\Delta \le 1$ *be fixed. Then the following bound holds:*

$$\begin{aligned}\mathbb{E}(\|\widetilde{f}_m - f\|^2) \quad &\le \quad \|f-f_m\|^2 + \frac{c}{n\Delta}\int_{-\pi m}^{\pi m} |f^*(u)|^2|u| \int_0^u \frac{1}{|f^*(v)|^2}\left(1+\left|\frac{(f^*)'(v)}{f^*(v)}\right|^2\right) dvdu \\ &\quad + \frac{c(p)}{(n\Delta)^p}\int_{-\pi m}^{\pi m} |u|^{2p-1}\int_0^u \frac{1}{|f^*(v)|^{2p}}\left(1+\left|\frac{(f^*)'(v)}{f^*(v)}\right|^{2p}\right) dvdu \end{aligned} \tag{20}$$

where $c, c(p)$ *are constants depending on* p, K_0*, the moments of* Λ *and* ξ_i *up to order* $2p$.
If moreover [B] (iii) *holds, we have*

$$\begin{aligned}\mathbb{E}(\|\widetilde{f}_m - f\|^2) \quad &\le \quad \|f-f_m\|^2 + \frac{c_1}{n\Delta}\int_{-\pi m}^{\pi m} |f^*(u)|^2 \left(\int_0^{|u|}\frac{dv}{|f^*(v)|^2} + \left(\int_0^{|u|}\frac{|(f^*)'(v)|}{|f^*(v)|^2}dv\right)^2\right) du \\ &\quad + \frac{c_2}{(n\Delta)^p}\int_{-\pi m}^{\pi m}\left(1+\left(\int_0^{|u|}\left|\frac{(f^*)'(v)}{f^*(v)}\right|^2 dv\right)^p\right)\left(\int_0^{|u|}\frac{dv}{|f^*(v)|^2}\right)^p du \\ &\quad + \frac{c_3}{(n\Delta)^{2p-1}}\int_{-\pi m}^{\pi m}\left(\int_0^{|u|}\frac{dv}{|f^*(v)|}\right)^{2p} du \end{aligned} \tag{21}$$

where the constants c_i, $i=1,2,3$ *depend on* $\|G'_\Delta\|_1$, K_0 *and the moments of* Λ *and* ξ_i *up to order* $2p$

The bounds are specific to our problem since the unknown function appears both in bias and variance terms.

3.2. **Rate of the estimator.** We study the resulting rate for the estimator on different examples.
• Gamma distribution. Let $f \sim \Gamma(\alpha, 1)$. Then $f^*(u) = (1-iu)^{-\alpha}$ and

$$(f^*)'(u)/f^*(u) = -\frac{i\alpha}{1-iu}.$$

Note that Assumption **[B]**(iii) is fulfilled.

Then $\|f - f_m\|^2 = O(m^{-2\alpha+1})$ so that $\alpha > 1/2$ is required for consistency of the estimator. For the variance terms, using the bound (20), we have

$$\mathbb{V}_1 := \int_{-\pi m}^{\pi m} |f^*(u)|^2 \left(\int_0^{|u|} \frac{dv}{|f^*(v)|^2} + \left(\int_0^{|u|} \frac{|(f^*)'(v)|}{|f^*(v)|^2} dv \right)^2 \right) du = O(m^2),$$

$$\mathbb{V}_2 := \int_{-\pi m}^{\pi m} \left(1 + \left(\int_0^{|u|} \left| \frac{(f^*)'(v)}{f^*(v)} \right|^2 dv \right)^p \right) \left(\int_0^{|u|} \frac{dv}{|f^*(v)|^2} \right)^p du = O(m^{(2\alpha+1)p+1}),$$

and

$$\mathbb{V}_3 := \int_{-\pi m}^{\pi m} \left(\int_0^{|u|} \frac{dv}{|f^*(v)|} \right)^{2p} du = O(m^{(2\alpha+1)p+1}).$$

Optimizing the bias and $\mathbb{V}_1/(n\Delta)$ yields $m_{opt,1} \asymp (n\Delta)^{1/(2\alpha+1)}$ and a rate $O((n\Delta)^{-(2\alpha-1)/(2\alpha+1)})$. Optimizing the bias and $\mathbb{V}_2/(n\Delta)^p$ yields $m_{opt,2} \asymp (n\Delta)^{1/(1+2\alpha(1+1/p))}$ and a rate

$$O((n\Delta)^{-(2\alpha-1)/(2\alpha+1+2\alpha/p)}).$$

Optimizing the bias and $\mathbb{V}_3/(n\Delta)^{2p-1}$ yields $m_{opt,3} \asymp (n\Delta)^{(2p-1)/(2p(\alpha+1)+2\alpha)}$ and a rate

$$O((n\Delta)^{-(2\alpha-1)(2p-1)/(2p(\alpha+1)+2\alpha)}).$$

For $p \geq 2$ the rate is of order $(n\Delta)^{-(2\alpha-1)/(2\alpha+1+\frac{2\alpha}{p})}$, which is close to $(n\Delta)^{-(2\alpha-1)/(2\alpha+1)}$ for large p. Thus, as p can be as large as possible, $\mathbb{V}_1$ and $\mathbb{V}_2$ get comparable.

• Gaussian distribution. Let us consider $f^*(u) = e^{-u^2/2}$, $(f^*)'(u)/f^*(u) = -u$. Assumption **[B]**(iii) is fulfilled. We use Lemma 6.2 recalled in the Appendix. Then $\|f - f_m\|^2 = O(m^{-1}e^{-(\pi m)^2})$, $\mathbb{V}_1 = O(m)$, $\mathbb{V}_2 = O(m^{2p-1}e^{p(\pi m)^2})$ and $\mathbb{V}_3 = O(m^{-2p-1}e^{p(\pi m)^2})$. We choose

$$(\pi m_{opt})^2 = \frac{p}{p+1} \log(n\Delta) - \frac{2p}{p+1} \log(\log(n\Delta)),$$

and get the rate

$$(n\Delta)^{-\frac{p}{p+1}} (\log n\Delta)^{\frac{p-1}{p+1}}.$$

Note that optimizing the bias and $\mathbb{V}_1/(n\Delta)$ yields the rate $\sqrt{\log(n\Delta)}/(n\Delta)$. Here again, for large p, the two terms $\mathbb{V}_1$ and $\mathbb{V}_2$ are comparable.

3.3. **Cut off selection.** As for the previous methods, we need to propose a data-driven selection of the cut off. As above, the bias is estimated up to a constant by $-\|\widetilde{f}_m\|^2$. The penalty is built by estimating the variance term of the risk bound. Here, we have three terms and we choose to estimate only the first one. So we set

$$\widetilde{m} = \arg\min_m (-\|\widetilde{f}_m\|^2 + \mathrm{qen}(m))$$

with

$$\mathrm{qen}(m) = \frac{\kappa \log^{\mathsf{a}}(n\Delta)}{n\Delta} \int_{-\pi m}^{\pi m} |\widetilde{f}^*(u)|^2 \left(\hat{s}_Y \int_0^{|u|} \frac{dv}{|\widetilde{f}^*(v)|^2} + \left(\int_0^{|u|} \frac{|\widetilde{\psi}(v)|}{|\widetilde{f}^*(v)|} dv \right)^2 \right) du,$$

where $\hat{s}_Y = n^{-1} \sum_{j=1}^n Y_j^2(\Delta)$ is an estimate of $\mathbb{E}(Y^2(\Delta))$ which appears when proving the risk bound.

We provide no theoretical result concerning this penalization criterion. As illustrated by the examples, estimating the first variance term $\mathbb{V}_1$ seems enough.

The constant κ is calibrated on preliminary simulations. The term $\log^{\mathsf{a}}(n\Delta)$ allows to stabilize the value of κ when $n\Delta$ varies and appears usually in adaptive deconvolution.

3.4. **Naive procedure.** A simple procedure is available for estimating f based on the joint observation $(N_j(\Lambda_j\Delta), Y_j(\Delta))_{1\leq j\leq n}$. This procedure is compared to the above nonparametric strategy on simulations.

Note that

$$\mathbb{P}(N_j(\Lambda_j\Delta) = 1) := \alpha_1(\nu, \Delta) = \Delta \int_0^{+\infty} e^{-\lambda\Delta}\lambda\nu(d\lambda) > 0$$

and that the conditional distribution of $Y_j(\Delta)$ given $N_j(\Lambda_j\Delta) = 1$ is identical to the distribution of ξ_1^j. Hence, let us set:

$$\frac{1}{\tilde{\alpha}_1} = \frac{1}{\hat{\alpha}_1}\mathbf{1}_{\hat{\alpha}_1 \geq k\sqrt{\Delta/n}}, \quad \hat{\alpha}_1 = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{(N_j(\Lambda_j\Delta)=1)}, \tag{22}$$

and

$$\check{f}_m(x) = \frac{1}{2\pi\tilde{\alpha}_1}\int_{-\pi m}^{\pi m} e^{-iux}\left(\frac{1}{n}\sum_{i=1}^{n} e^{iuY_j(\Delta)}\mathbf{1}_{(N_j(\Lambda_j\Delta)=1)}\right)du. \tag{23}$$

The following property holds.

Proposition 3.2. *Assume* **(H0)***,* $\mathbb{E}(\Lambda) < +\infty$*,* $\mathbb{E}(\Lambda e^{-\Lambda\Delta}) \geq k_0$ *and* $n\Delta \geq 1 \vee \frac{4k^2}{k_0^2}$*. Then*

$$\mathbb{E}(\|\check{f}_m - f\|^2) \leq \|f - f_m\|^2 + \frac{4m}{n\alpha_1(\nu,\Delta)}\left(1 + \frac{2k^2\Delta}{\alpha_1(\nu,\Delta)}\frac{1}{n\alpha_1(\nu,\Delta)}\right)$$

where f_m *is such that* $f_m^* = f^*\mathbf{1}_{[-\pi m, \pi, m]}$

Note that $\alpha_1(\nu,\Delta) = \Delta(\mathbb{E}(\Lambda) + o(1))$. The variance term is of order $O(m/(n\Delta))$. We propose the following adaptive choice for m:

$$\check{m} = \arg\min_{m \leq n\Delta}\left(-\|\check{f}_m\|^2 + \frac{\kappa}{n\tilde{\alpha}_1}\right).$$

The proof of Proposition 3.2 follows the same lines as the analogous Proposition 2.2 and is omitted. The proof of adaptiveness of $\check{f}_{\check{m}}$ is also omitted.
The interest of the naive estimator is obviously its simplicity. However, it strongly depends on the observed value $\hat{\alpha}_1$ as the number of observations taken for the estimation is $n\hat{\alpha}_1$. If this value is too small, the estimator performs poorly.

4. Illustrations of the methods

In this section, we illustrate the estimators with data driven cut off on simulated data for Λ. Tables 1, 2, 3, 4 correspond to f a Gaussian $\mathcal{N}(0,3)$ and Λ either an exponential distribution with mean 1, or a uniform density on $[5,6]$, a translated exponential $\mathcal{E}(2) + 5$, a translated $\mathcal{B}eta$ distribution $\mathcal{B}eta(2,2) + 5$. Figure 1 corresponds to a jump density mixture of Gaussian $0.4\mathcal{N}(-2,1) + 0.6\mathcal{N}(3,1)$ and Λ an exponential $\mathcal{E}(1)$. Figure 2 plots a Gumbel jump distribution with c.d.f. $F(x) = \exp(-\exp(-x)), x > 0$ and Λ either $\mathcal{E}(1)$ or $\mathcal{U}([1,2])$. The truncation constant k is alway taken equal to 0.5, except in (22) where $k = 0$.

Δ	0.01	0.1	0.5	0.9	1	2
n	20000	2000	400	220	200	100
$\overline{n}_{\neq 0}$	200	181	133	104	99	66
L_2 Risk	0.14	$6.2 \cdot 10^{-3}$	$3.9 \cdot 10^{-3}$	$5.3 \cdot 10^{-3}$	$5.6 \cdot 10^{-3}$	$2.1 \cdot 10^{-2}$
	$(5.9 \cdot 10^{-3})$	$(2.8 \cdot 10^{-3})$	$(2.6 \cdot 10^{-3})$	$(4.8 \cdot 10^{-3})$	$(5.7 \cdot 10^{-3})$	(0.06)
$\overline{\hat{m}}$	0.01	0.25	0.29	0.29	0.29	0.29
n	100000	10000	2000	1110	1000	500
$\overline{n}_{\neq 0}$	990	908	666	526	499	333
L_2 Risk	0.014	$1.1 \cdot 10^{-3}$	$9.1 \cdot 10^{-4}$	$1.4 \cdot 10^{-3}$	$1.5 \cdot 10^{-3}$	$3.5 \cdot 10^{-3}$
	(0.002)	$(5.0 \cdot 10^{-4})$	$(5.8 \cdot 10^{-4})$	$(1.6 \cdot 10^{-3})$	$(1.7 \cdot 10^{-3})$	$(5.0 \cdot 10^{-3})$
$\overline{\hat{m}}$	0.21	0.34	0.37	0.36	0.36	0.33
n	-	50000	10000	5550	5000	2500
$\overline{n}_{\neq 0}$	-	4545	3333	2969	2500	1666
L_2 Risk	-	$2.4 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$	$3.7 \cdot 10^{-4}$	$9.2 \cdot 10^{-4}$
	-	$(9.5 \cdot 10^{-5})$	$(1.3 \cdot 10^{-4})$	$(2.3 \cdot 10^{-4})$	$(2.9 \cdot 10^{-4})$	$(1.1 \cdot 10^{-3})$
$\overline{\hat{m}}$	-	0.24	0.43	0.42	0.42	0.39

TABLE 1. Mean of the $\mathbb{L}_2$-risks for the semi-parametric method 1 with $\Lambda \sim \mathcal{E}(1)$ (standard deviation in parenthesis) and f is $\mathcal{N}(0,3)$; $\overline{n}_{\neq 0}$ is the mean of nonzero data; $\overline{\hat{m}}$ is the mean of selected $\hat{m}$'s.

All methods require the calibration of the constant κ in penalties and the exponent $\mathtt{a}$ in method 3. After preliminary experiments, κ is taken equal to 0.21 in method 1 (first semi-parametric method), to 5 in method 2 (second semi-parametric method), 0.15 in method 3 with $\alpha = 1.75$ (pure nonparametric method), and to 5 in the naive method. The cut off m is selected among 200 equispaced values between 0.01 and 2.

Δ	0.01	0.1	0.5	0.9	1
n	20000	2000	400	220	200
$\overline{n}_{\neq 0}$	197	183	132	104	100
K	1	1	3	25	50
L_2 Risk	$1.6 \cdot 10^{-3}$	$1.9 \cdot 10^{-3}$	$3.8 \cdot 10^{-3}$	$2.6 \cdot 10^{-2}$	$1.5 \cdot 10^{16}$
	$(1.3 \cdot 10^{-3})$	$(1.3 \cdot 10^{-3})$	$(2.9 \cdot 10^{-3})$	$(4.7 \cdot 10^{-2})$	$(5.0 \cdot 10^{17})$
$\overline{\hat{m}_K}$	0.36	0.35	0.30	0.25	0.24
n	100000	10000	2000	1110	1000
$\overline{n}_{\neq 0}$	992	910	667	525	499
K	1	1	4	32	50
L_2 Risk	$4.4 \cdot 10^{-4}$	$4.6 \cdot 10^{-4}$	$9.4 \cdot 10^{-4}$	$5.4 \cdot 10^{-3}$	$2.3 \cdot 10^{3}$
	$(3.5 \cdot 10^{-2})$	$(3.1 \cdot 10^{-4})$	$(6.6 \cdot 10^{-4})$	$(3.5 \cdot 10^{-2})$	$(3.1 \cdot 10^{4})$
$\overline{\hat{m}_K}$	0.44	0.42	0.38	0.35	0.34

TABLE 2. Mean of the $\mathbb{L}_2$-risks for the semi-parametric method 2 with $\Lambda \sim \mathcal{E}(1)$ (standard deviation in parenthesis) and f is $\mathcal{N}(0,3)$; $\overline{n}_{\neq 0}$ is the mean of nonzero data; $\overline{\hat{m}}$ is the mean of selected $\hat{m}_K$'s.

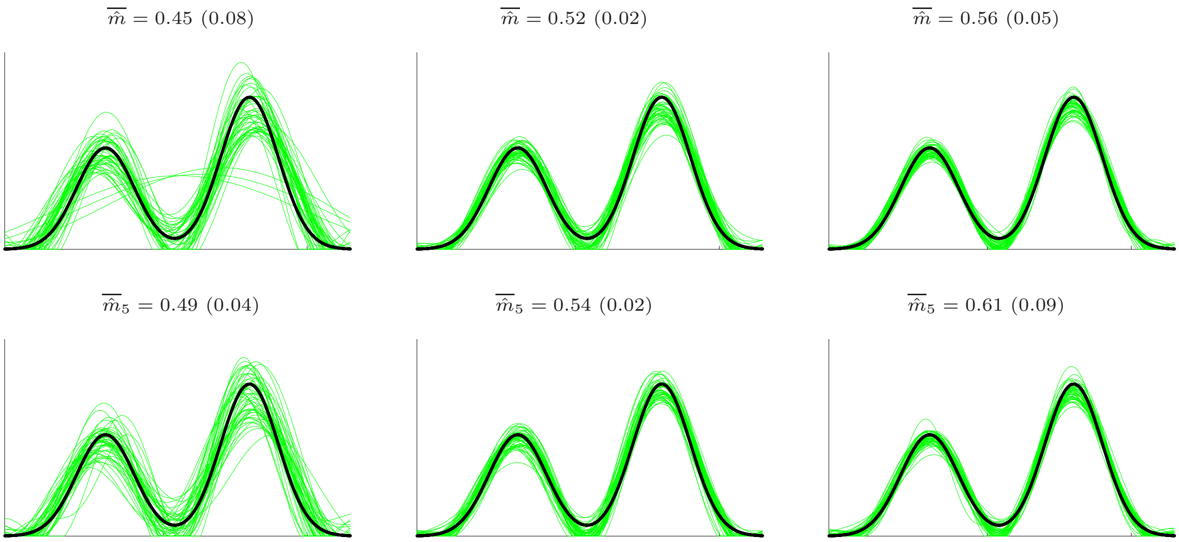

FIGURE 1. Estimation of f for a Gaussian mixture $0.4\mathcal{N}(-2,1)+0.6\mathcal{N}(3.1)$ for $n = 500$ (first column) $n = 2000$ (second column) and $n = 5000$ (third column) with the semiparametric method 1 (first line) and the semi parametric method 2 (second line, $K = 5$) for $\Delta = 1/2$. True density (bold black line) and 50 estimated curves (green lines).

We give $\mathbb{L}_2$-risks for $f \sim \mathcal{N}(0,3)$ and graphical illustrations for the two other examples of jump densities. To compute $\mathbb{L}_2$-risks, we use 1000 repetitions for $n\Delta = 200, 1000$ and 2500 repetitions for $n\Delta = 5000$. Note that because of computational limitations, we only make 100 repetitions when $\Delta = 0.01$ (first columns of Tables 1, 2).

Tables 1, 2 allow to compare the two semiparametric methods for the Gaussian jump density and Λ exponential $\mathcal{E}(1)$. Method 1 works for fixed values of Δ ($\Delta = 1, 2$), but also for small values (0.1 to 0.9). However, when Δ gets too small (0.01), the risk increases. On the other hand method 2, completely fails for $\Delta = 1$, as predicted by the theory ($A = +\infty$ in the risk bound of Proposition 2.5 when $\mu\Delta = 1$); for $\Delta = 0.5, 0.9$, methods 1 and 2 have comparable risks while for $\Delta \leq 0.1$, method 2 is better. The value $\bar{n}_{\neq 0}$ of non zero data represents the number of data really used for estimation. The cut off values are rather small and stable (standard deviations are of order 10^{-2}). The value K is taken of order $\sup(1, K_0)$ for K_0 defined in formula (17). In Figure 1, 50 estimated curves of a Gaussian mixture by methods 1 and 2 are plotted, for different sample sizes $n = 500, 2000, 5000$. The two methods distinguish well the two modes and are improved as n increases.

Table 3 shows that the first semi-parametric, the pure nonparametric and the naive procedures are comparable, with good results even for small values of n. The naive method performs surprisingly well and is stable. In Table 4, we change the distribution of Λ and therefore we show no results for method 1, since it does not work in that case, neither in theory nor in practice. The chosen distributions for Λ make the naive method perform worse than method 3. For $n = 200$, the risk of method 3 is twice better, for $n = 500$, it is three times better and for $n = 1000$, four times better. For larger n, the methods become equivalent. The naive method

n	Method 1	Method 3	Naive
200	$5.6 \cdot 10^{-3}$ $(5.7 \cdot 10^{-3})$	$4.9 \cdot 10^{-3}$ $(6.0 \cdot 10^{-3})$	$5.9 \cdot 10^{-3}$ $(4.5 \cdot 10^{-3})$
500	$2.7 \cdot 10^{-3}$ $(2.5 \cdot 10^{-3})$	$3.6 \cdot 10^{-3}$ $(1.2 \cdot 10^{-3})$	$2.6 \cdot 10^{-3}$ $(2.2 \cdot 10^{-3})$
1000	$1.5 \cdot 10^{-3}$ $(1.7 \cdot 10^{-3})$	$2.4 \cdot 10^{-3}$ $(5.2 \cdot 10^{-3})$	$1.3 \cdot 10^{-3}$ $(9.2 \cdot 10^{-4})$
2000	$7.9 \cdot 10^{-4}$ $(5.7 \cdot 10^{-4})$	$2.0 \cdot 10^{-3}$ $(9.7 \cdot 10^{-3})$	$7.2 \cdot 10^{-4}$ $(5.1 \cdot 10^{-4})$
5000	$3.7 \cdot 10^{-4}$ $(2.9 \cdot 10^{-4})$	$7.4 \cdot 10^{-4}$ $(9.4 \cdot 10^{-4})$	$3.1 \cdot 10^{-4}$ $(2.2 \cdot 10^{-4})$

TABLE 3. Mean of the $\mathbb{L}_2$-risks for the semi-parametric method 1, the nonparametric method 3 and the naive method (see section 3.4). $\Lambda \sim \mathcal{E}(1)$ and f is $\mathcal{N}(0,3)$; standard deviation in parenthesis.

fails here because the number $n\hat{\alpha}_1$ (see (22)) is too small. In Figure 2, 50 estimated curves of a Gumbel distribution by methods 1 and 3 are plotted, for different sample sizes $n = 500, 2000$, for Λ an exponential $\mathcal{E}(1)$ and a uniform $\mathcal{U}([1,2])$. Columns 1 and 2 allow to compare methods 1 and 3 when Λ is exponential. Method 3 has good performances without estimating the distribution of Λ. In all cases, the values of m are small and stable.

Λ	n	Method 3	Naive
$\mathcal{U}([5,6])$	200	$2.8 \cdot 10^{-2}$ $(6.6 \cdot 10^{-3})$	$5.6 \cdot 10^{-2}$ $(1.5 \cdot 10^{-2})$
	500	$9.0 \cdot 10^{-3}$ $(3.8 \cdot 10^{-3})$	$2.8 \cdot 10^{-2}$ $(1.8 \cdot 10^{-2})$
	1000	$3.3 \cdot 10^{-3}$ $(2.2 \cdot 10^{-3})$	$1.3 \cdot 10^{-2}$ $(9.5 \cdot 10^{-3})$
$\mathcal{E}(2)+5$	200	$2.8 \cdot 10^{-2}$ $(6.8 \cdot 10^{-3})$	$5.6 \cdot 10^{-2}$ $(1.7 \cdot 10^{-2})$
	500	$9.0 \cdot 10^{-3}$ $(3.7 \cdot 10^{-3})$	$2.7 \cdot 10^{-2}$ $(1.6 \cdot 10^{-2})$
	1000	$3.4 \cdot 10^{-3}$ $(2.3 \cdot 10^{-3})$	$1.3 \cdot 10^{-2}$ $(8.7 \cdot 10^{-3})$
$\mathcal{B}eta(2,2)+5$	200	$2.7 \cdot 10^{-2}$ $(6.5 \cdot 10^{-3})$	$5.6 \cdot 10^{-2}$ $(1.4 \cdot 10^{-2})$
	500	$9.2 \cdot 10^{-3}$ $(4.0 \cdot 10^{-3})$	$2.9 \cdot 10^{-2}$ $(1.7 \cdot 10^{-2})$
	1000	$3.3 \cdot 10^{-3}$ $(2.2 \cdot 10^{-3})$	$1.3 \cdot 10^{-2}$ $(8.9 \cdot 10^{-3})$

TABLE 4. Mean of the $\mathbb{L}_2$-risks for the nonparametric method 3 and the naive method (see section 3.4); standard deviation in parenthesis. $\Lambda \sim \mathcal{U}([5,6])$, $\mathcal{E}(2)+5$, $\mathcal{B}eta(2,2)+5$ and f is $\mathcal{N}(0,3)$.

5. Appendix: Proofs

5.1. **Proof of Proposition 2.1.** By the Rosenthal Inequality, we get

$$\mathbb{E}\left(|\hat{q}_\Delta - q_\Delta|^{2p}\right) \le \frac{\Lambda(2p)}{n^{2p}}\left(nq_\Delta + (nq_\Delta(1-q_\Delta))^p\right).$$

Now, $nq_\Delta = n\Delta\mu/(1+\mu\Delta) \le n\Delta\mu_1$ gives, under $n\Delta \ge 1$

$$\mathbb{E}\left(|\hat{q}_\Delta - q_\Delta|^{2p}\right) \le \frac{C(2p)\Delta^p}{n^p}(\mu_1 + \mu_1^p)$$

and thus the first bound. Now we write that

$$\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta} = \left(\frac{1}{\hat{q}_\Delta} - \frac{1}{q_\Delta}\right)\mathbf{I}_{\hat{q}_\Delta > q_{0,\Delta}/2} - \frac{1}{q_\Delta}\mathbf{I}_{\hat{q}_\Delta \le q_{0,\Delta}/2}.$$

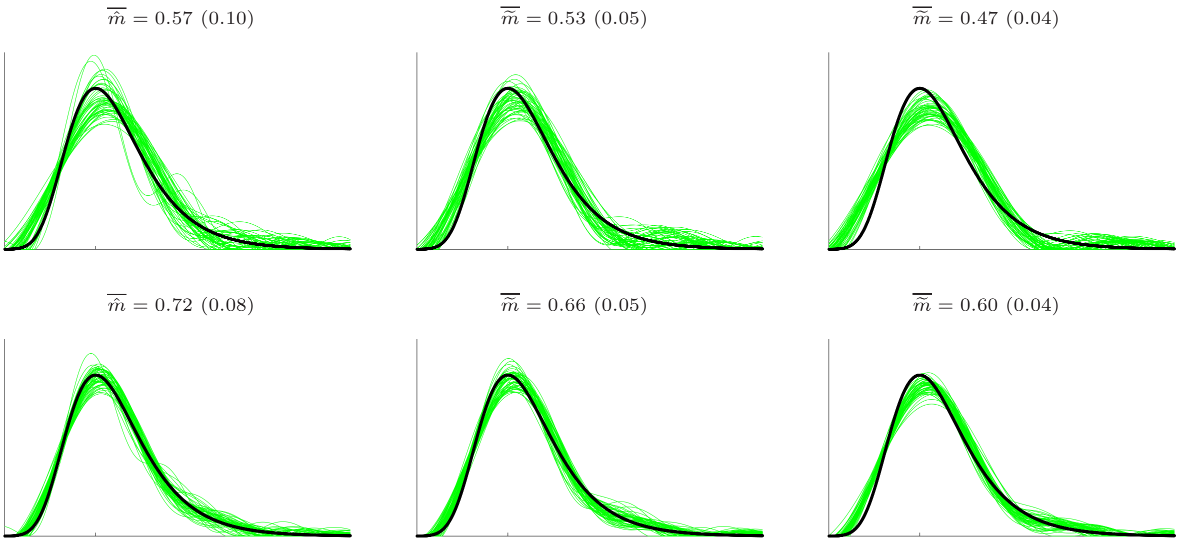

FIGURE 2. Estimation of f for a Gumbel distribution, for $n = 500$ (first line) $n = 2000$ (second line) with method 1 (first column) and method 3 (second and third column) for $\Delta = 1$. In the first two columns Λ is $\mathcal{E}(1)$ and in the third Λ is $\mathcal{U}([1, 2])$. True density (bold black line) and 50 estimated curves (green lines).

We get

$$\begin{aligned}
\mathbb{E}\left(\left|\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta}\right|^{2p}\right) &\leq \frac{\mathbb{E}(|\hat{q}_\Delta - q_\Delta|^{2p})}{(q_{0,\Delta}^2/2)^{2p}} + \frac{1}{q_{0,\Delta}^{2p}}\mathbb{P}(|\hat{q}_\Delta - q_\Delta| > q_{0,\Delta}/2) \\
&\leq \frac{2^{2p}}{q_{0,\Delta}^{4p}}\mathbb{E}(|\hat{q}_\Delta - q_\Delta|^{2p}) + \frac{1}{q_{0,\Delta}^{2p}}\frac{\mathbb{E}(|\hat{q}_\Delta - q_\Delta|^{2p}}{(q_{0,\Delta}/2)^{2p}} \\
&\leq \frac{2^{2p+1}}{q_{0,\Delta}^{4p}}\mathbb{E}(|\hat{q}_\Delta - q_\Delta|^{2p}) \leq \frac{C'(p,\Delta)}{(n\Delta^3)^p},
\end{aligned}$$

with $C'(p,\Delta) = 2^{2p+1}C(2p)(1+\mu_0\Delta)^{4p}(\mu_1 + \mu_1^p)/\mu_0^{4p}$.

For $\tilde{\mu}$, the decomposition is the following

$$\begin{aligned}
\tilde{\mu} - \mu &= \frac{1}{\Delta}\left\{\frac{\hat{q}_\Delta - q_\Delta}{1 - \hat{q}_\Delta}\mathbb{1}_{1-\hat{q}_\Delta > (1-q_{1,\Delta})/2} + q_\Delta\left(\frac{1}{\hat{q}_\Delta} - \frac{1}{q_\Delta}\right)\mathbb{1}_{1-\hat{q}_\Delta > (1-q_{1,\Delta})/2}\right. \\
&\qquad \left. - \frac{q_\Delta}{1-q_\Delta}\mathbb{1}_{1-\hat{q}_\Delta \leq (1-q_{1,\Delta})/2}\right\}.
\end{aligned}$$

This yields

$$\begin{aligned}
\mathbb{E}(|\tilde{\mu} - \mu|^{2p}) &\leq \frac{\mathbb{E}(|\hat{q}_\Delta - q_\Delta|^{2p})}{\Delta^{2p}}\left\{\frac{2^{2p-1}}{((1-q_{1,\Delta})/2)^{2p}} + \frac{2^{2p-1}q_\Delta^{2p}}{(1-q_\Delta)^{2p}[(1-q_{1,\Delta})/2]^{2p}}\right. \\
&\qquad \left. + \left(\frac{q_\Delta}{1-q_\Delta}\right)^{2p}\frac{1}{((1-q_{1,\Delta})/2)^{2p}}\right\} \\
&\leq \frac{C''(p,\Delta)}{(n\Delta)^p},
\end{aligned}$$

with $C''(p,\Delta) = 2^{2p-1}(1+2(\mu_1\Delta)^{2p})C(2p)(\mu_1+\mu_1^{2p})/(1-q_{1,\Delta})/2$. $\square$

5.2. **Proof of Proposition 2.2.** We use the fact that $|\phi_\Delta(u)|^{-1} \leq 1+2\mu\Delta$, $q_\Delta^{-1} \leq 1+1/(\mu_0\Delta)$, $\|f^*\|^2 = 2\pi\|f\|^2$, Proposition 2.1 and the following Lemma, proved in Section 5.3.

Lemma 5.1.

$$\forall u \in \mathbb{R}, \quad \mathbb{E}(|\frac{1}{\widetilde{\phi_\Delta}(u)} - \frac{1}{\phi_\Delta(u)}|^{2p}) \leq c_p \left(\frac{1}{|\phi_\Delta(u)|^{2p}} \wedge \frac{n^{-p}}{|\phi_\Delta(u)|^{4p}}\right).$$

In our setting, only the second term of the bound is useful.

Moreover Rosenthal's inequality implies for $p \geq 1$,

$$\mathbb{E}\left(|\hat{Q}_\Delta(u) - Q_\Delta(u)|^{2p}\right) \leq \frac{C(2p)}{n^{2p}}(nq_\Delta + (nq_\Delta)^p).$$

Consequently,

$$\mathbb{E}\left(|\hat{Q}_\Delta(u) - Q_\Delta(u)|^{2p}\right) \leq c_p(\mu_1+\mu_1^p)\left(\frac{\Delta}{n}\right)^p. \tag{24}$$

Then we have the decomposition

$$\hat{f}^*(u) - f^*(u) = T_0(u) + \sum_{i=1}^{6} T_i(u), \tag{25}$$

with

$$T_0(u) = \frac{\hat{Q}_\Delta(u) - Q_\Delta(u)}{q_\Delta\phi_\Delta(u)}$$

$$T_1(u) = \Big(\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta}\Big)\frac{Q_\Delta(u)}{\phi_\Delta(u)}, \quad T_2(u) = (\hat{Q}_\Delta(u) - Q_\Delta(u))\Big(\frac{1}{\widetilde{\phi_\Delta}(u)} - \frac{1}{\phi_\Delta(u)}\Big)\Big(\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta}\Big)$$

$$T_3(u) = Q_\Delta(u)\Big(\frac{1}{\widetilde{\phi_\Delta}(u)} - \frac{1}{\phi_\Delta(u)}\Big), \quad T_4(u) = \frac{\hat{Q}_\Delta(u) - Q_\Delta(u)}{\phi_\Delta(u)}\Big(\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta}\Big)$$

$$T_5(u) = \frac{Q_\Delta(u)}{q_\Delta}\Big(\frac{1}{\widetilde{\phi_\Delta}(u)} - \frac{1}{\phi_\Delta(u)}\Big), \quad T_6(u) = \frac{\hat{Q}_\Delta(u) - Q_\Delta(u)}{q_\Delta}\Big(\frac{1}{\widetilde{\phi_\Delta}(u)} - \frac{1}{\phi_\Delta(u)}\Big),$$

Then

$$\int_{-\pi m}^{\pi m} |\hat{f}^*(u) - f^*(u)|^2 du \leq 2\int_{-\pi m}^{\pi m} |T_0(u)|^2 du + 12\sum_{i=1}^{6}\int_{-\pi m}^{\pi m} |T_i(u)|^2 du$$

and first

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_0(u)|^2 du\right) \leq \frac{1}{nq_\Delta}\int_{-\pi m}^{\pi m} \frac{du}{|\phi_\Delta(u)|^2}.$$

For the following bounds, we use constants c, c' which may change from line to line but depend neither on n nor on Δ. We have

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_1(u)|^2 du\right) = \int_{-\pi m}^{\pi m} |f^*(u)|^2 du\, \mathbb{E}\left(q_\Delta^2\Big|\frac{1}{\tilde{q}_\Delta} - \frac{1}{q_\Delta}\Big|^2\right) \leq \frac{c}{n\Delta}2\pi\|f\|^2,$$

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_2(u)|^2 du\right) \leq \frac{c}{n^3\Delta^2}\int_{-\pi m}^{\pi m} \frac{du}{|\phi_\Delta(u)|^4} \leq \frac{c'}{n^2\Delta},$$

as $m \leq n\Delta$. Next

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_3(u)|^2 du\right) \leq \frac{c}{n^2\Delta}\int_{-\pi m}^{\pi m} \frac{du}{|\phi_\Delta(u)|^4} \leq \frac{c'}{n},$$

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_4(u)|^2 du\right) \leq \frac{c}{n^2\Delta^2}\int_{-\pi m}^{\pi m} \frac{du}{|\phi_\Delta(u)|^2} \leq \frac{c'}{n\Delta},$$

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_5(u)|^2 du\right) \leq \frac{c}{n}\int_{-\pi m}^{\pi m} \frac{|f^*(u)|^2}{|\phi_\Delta(u)|^2} du \leq \frac{c'\|f\|^2}{n},$$

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} |T_6(u)|^2 du\right) \leq \frac{c}{n^2\Delta}\int_{-\pi m}^{\pi m} \frac{du}{|\phi_\Delta(u)|^4} \leq \frac{c'}{n}.$$

Gathering all the terms gives the result of Proposition 2.2.□

5.3. **Proof of Lemma 5.1.** Below, c is a constant which may change from line to line.

Let $B_n = (|\widehat{\phi}_\Delta(u)| \geq k/\sqrt{n})$. We write

$$\frac{1}{\widetilde{\phi_\Delta}(u)} - \frac{1}{\phi_\Delta(u)} = A_1 + A_2$$

with

$$A_1 = \left(\frac{1}{\widehat{\phi}_\Delta(u)} - \frac{1}{\phi_\Delta(u)}\right) \mathbf{1}_{B_n}, \quad A_2 = -\frac{1}{\phi_\Delta(u)} \mathbf{1}_{B_n^c}.$$

We have:

$$|A_1| \leq \frac{\sqrt{n}|\widehat{\phi}_\Delta(u) - \phi_\Delta(u)|}{k|\phi_\Delta(u)|}.$$

By the Rosenthal inequality, for all $q \geq 1$, $\mathbb{E}|\widehat{\phi}_\Delta(u) - \phi_\Delta(u)|^{2q} \leq 3/n^q$. Therefore, $\mathbb{E}|A_1|^{2p} \leq (3/k)|\phi_\Delta(u)|^{2p}$. Obviously, $\mathbb{E}|A_2|^{2p} \leq 1/|\phi_\Delta(u)|^{2p}$. Now, we proceed to find the other bound. We have $A_1 = A_1' + A_1''$ with

$$A_1' = \frac{(\widehat{\phi}_\Delta(u) - \phi_\Delta(u))^2}{\phi_\Delta^2(u)\widehat{\phi}_\Delta(u)} \mathbf{1}_{B_n}, \quad A_1'' = \frac{\phi_\Delta(u) - \widehat{\phi}_\Delta(u)}{\phi_\Delta^2(u)} \mathbf{1}_{B_n}.$$

We have:

$$\mathbb{E}|A_1'|^{2p} \leq \frac{1}{|\phi_\Delta(u)|^{4p}}\left(\frac{\sqrt{n}}{k}\right)^{2p}\frac{3}{n^{2p}} = \frac{3}{k^{2p}}\frac{n^{-p}}{|\phi_\Delta(u)|^{4p}}, \quad \mathbb{E}|A_1''|^{2p} \leq 3\frac{n^{-p}}{|\phi_\Delta(u)|^{4p}}.$$

Therefore, we also have $\mathbb{E}|A_1|^{2p} \leq cn^{-p}/\phi_\Delta(u)|^{4p}$. Now, we study the second inequality for the term A_2. First note that

$$B_n^c \subset (|\widehat{\phi}_\Delta(u) - \phi_\Delta(u)| \geq |\phi_\Delta(u)| - k/\sqrt{n}).$$

Moreover, $|\phi_\Delta(u)| - k/\sqrt{n} > |\phi_\Delta(u)|/2 \iff |\phi_\Delta(u)| > 2k/\sqrt{n}$. Thus,

$$\begin{aligned}
\mathbb{P}(B_n^c) &\leq \mathbb{P}\left(|\widehat{\phi}_\Delta(u) - \phi_\Delta(u)| \geq \frac{|\phi_\Delta(u)|}{2}\right) + \mathbf{1}_{(|\phi_\Delta(u)| \leq 2k/\sqrt{n})} \\
&\leq \left(\frac{2}{|\phi_\Delta(u)|}\right)^{2p} \mathbb{E}|\widehat{\phi}_\Delta(u) - \phi_\Delta(u)|^{2p} + \mathbf{1}_{(|\phi_\Delta(u)|^{-1} \geq \sqrt{n}/2k)} \\
&\leq c\left(\frac{2}{|\phi_\Delta(u)|}\right)^{2p} n^{-p} + \left(\frac{2k}{\sqrt{n}}\right)^{2p} |\phi_\Delta(u)|^{-2p}.
\end{aligned}$$

Thus $\mathbb{P}(B_n^c) \leq cn^{-p}/|\phi_\Delta(u)|^{2p}$. Finally, we also have:

$$\mathbb{E}|A_2|^{2p} \leq c\frac{n^{-p}}{|\phi_\Delta(u)|^{4p}}. \tag{26}$$

So the proof of Lemma 5.1 is complete. □

5.4. **Proof of Theorem 2.1.** Let

$$S_m = \{t \in \mathbb{L}^2(\mathbb{R}),\ t^* = t^* \mathbf{1}_{[-\pi m, \pi m]}\},$$

and consider the contrast

$$\gamma_n(t) = \|t\|^2 - \frac{2}{2\pi}\langle t^*, \hat{f}^*\rangle.$$

Clearly, $\hat{f}_m = \arg\min_{t\in S_m} \gamma_n(t)$ and $\gamma_n(\hat{f}_m) = -\|\hat{f}_m\|^2$. Moreover, we note that

$$\gamma_n(t) - \gamma_n(s) = \|t-f\|^2 - \|s-f\|^2 - \frac{2}{2\pi}\langle t^* - s^*, \hat{f}^* - f^*\rangle. \tag{27}$$

By definition of $\hat{m}$, we have

$$\gamma_n(\hat{f}_{\hat{m}}) + \widehat{\mathrm{pen}}(\hat{m}) \leq \gamma_n(f_m) + \widehat{\mathrm{pen}}(m).$$

This with (27) implies

$$\|\hat{f}_{\hat{m}} - f\|^2 \leq \|f - f_m\|^2 + \widehat{\mathrm{pen}}(m) + \frac{2}{2\pi}\langle \hat{f}^*_{\hat{m}} - f^*_m, \hat{f}^* - f^*\rangle - \widehat{\mathrm{pen}}(\hat{m}). \tag{28}$$

Writing that

$$\begin{aligned}
2\langle \hat{f}^*_{\hat{m}} - f^*_m, \hat{f}^* - f^*\rangle &\leq 2\|\hat{f}^*_{\hat{m}} - f^*_m\| \sup_{t\in S_{\hat{m}}+S_m, \|t\|=1} |\langle t^*, \hat{f}^* - f^*\rangle| \\
&\leq \frac{1}{4}\|\hat{f}^*_{\hat{m}} - f^*_m\|^2 + 4 \sup_{t\in S_{\hat{m}\vee m}, \|t\|=1} \langle t^*, \hat{f}^* - f^*\rangle^2 \\
&\leq \frac{1}{2}\|\hat{f}^*_{\hat{m}} - f^*\|^2 + \frac{1}{2}\|f^* - f^*_m\|^2 + 4 \sup_{t\in S_{\hat{m}\vee m}, \|t\|=1} \langle t^*, \hat{f}^* - f^*\rangle^2,
\end{aligned}$$

plugging this in (28) and gathering the terms, we get

$$\frac{1}{2}\|\hat{f}_{\hat{m}} - f\|^2 \leq \frac{3}{2}\|f - f_m\|^2 + \widehat{\mathrm{pen}}(m) + \frac{4}{2\pi} \sup_{t\in S_{\hat{m}\vee m}, \|t\|=1} \langle t^*, \hat{f}^* - f^*\rangle^2 - \widehat{\mathrm{pen}}(\hat{m}). \tag{29}$$

Now, we write the decomposition

$$\langle t^*, \hat{f}^* - f^*\rangle = \frac{1}{q_\Delta}\langle t^*, \frac{\hat{Q}_\Delta - Q_\Delta}{\phi_\Delta}\rangle + R(t)$$

where $R(t) = \sum_{i=1}^6 \langle t^*, T_i\rangle$ where the T_i's are defined by (25).

Clearly, the proof of Proposition 2.2, the Cauchy Schwarz inequality and $\|t^*\|^2 = 2\pi$ lead to

$$\mathbb{E}\left(\sup_{t\in S_{\hat{m}\vee m}, \|t\|=1} \left|R(t)\right|^2\right) \leq \frac{c}{n\Delta}.$$

Thus, we have to study the term

$$\sup_{t\in S_{\hat{m}\vee m}, \|t\|=1} \frac{1}{q_\Delta}\langle t^*, \frac{\hat{Q}_\Delta - Q_\Delta}{\phi_\Delta}\rangle,$$

for which, we can prove (see the proof in Section 5.5):

Lemma 5.2. *Assume that the Assumptions of Theorem 2.1 hold. Let*

$$p(m,\hat m)=\frac{3}{2\pi n q_\Delta}\int_{-\pi(m\vee\hat m)}^{\pi(m\vee\hat m)}\frac{du}{|\phi_\Delta(u)|^2},$$

we have

$$\mathbb{E}\left(\frac{1}{2\pi}\sup_{t\in S_{\hat m\vee m},\|t\|=1}\langle t^*,\frac{\hat Q_\Delta-Q_\Delta}{q_\Delta\phi_\Delta}\rangle^2-p(m,\hat m)\right)_+\leq\frac{c}{n\Delta}.$$

Let us define

$$\Omega=\left\{\left|\frac{1}{\tilde q_\Delta}-\frac{1}{q_\Delta}\right|\leq\frac{1}{2q_\Delta}\right\}.$$

and

$$\mathrm{pen}(m)=\frac{1}{2\pi n q_\Delta}\int_{-\pi m}^{\pi m}\frac{du}{|\phi_\Delta(u)|^2}.$$

We have $p(m,m')\leq 3\mathrm{pen}(m)+3\mathrm{pen}(m')$ and on Ω, we have, $\forall m\in\mathcal{M}_n$,

$$\begin{aligned}\mathbb{E}(\widehat{\mathrm{pen}}(m)\mathbf{1}_\Omega) &\leq \frac{3}{2}\frac{\kappa}{2\pi n q_\Delta}\mathbb{E}\left(\int_{-\pi m}^{\pi m}\frac{du}{|\widetilde{\phi_\Delta}(u)|^2}\right)\\ &\leq \frac{3\kappa}{2\pi n q_\Delta}\int_{-\pi m}^{\pi m}\frac{du}{|\phi_\Delta(u)|^2}+\frac{3\kappa}{2\pi n q_\Delta}\int_{-\pi m}^{\pi m}\mathbb{E}\left(\left|\frac{1}{\widetilde{\phi_\Delta}(u)}-\frac{1}{\phi_\Delta(u)}\right|^2\right)du\\ &\leq 3\kappa\ \mathrm{pen}(m)+\frac{3\kappa}{2\pi n q_\Delta}\frac{2\pi m(1+2\mu\Delta)^4}{n}\leq 3\kappa\ \mathrm{pen}(m)+\frac{c}{n}.\end{aligned}$$

Using (29) and Lemma 5.2, we derive

$$\begin{aligned}\mathbb{E}(\|\hat f_{\hat m}-f\|^2\mathbf{1}_\Omega) &\leq 3\|f-f_m\|^2+6\kappa\mathrm{pen}(m)+\mathbb{E}\left([16p(m,\hat m)-2\widehat{\mathrm{pen}}(\hat m)]\mathbf{1}_\Omega\right)+\frac{c}{n\Delta}\\ &\leq 3\|f-f_m\|^2+6(\kappa+8)\mathrm{pen}(m)+2\mathbb{E}\left([24\mathrm{pen}(\hat m)-\widehat{\mathrm{pen}}(\hat m)]\mathbf{1}_\Omega\right)+\frac{c}{n\Delta}.\end{aligned}\tag{30}$$

Now we note that

$$\begin{aligned}\mathbb{E}\left([16\mathrm{pen}(\hat m)-\widehat{\mathrm{pen}}(\hat m)]\mathbf{1}_\Omega\right) &= \mathbb{E}\left[\left(\frac{24}{2\pi n q_\Delta}\int_{-\pi\hat m}^{\pi\hat m}\frac{du}{|\phi_\Delta(u)|^2}-\frac{\kappa}{2\pi n\tilde q_\Delta}\int_{-\pi\hat m}^{\pi\hat m}\frac{du}{|\widetilde{\phi_\Delta}(u)|^2}\right)\mathbf{1}_\Omega\right]\\ &\leq \mathbb{E}\left[\left(\frac{96}{2\pi n\tilde q_\Delta}\int_{-\pi\hat m}^{\pi\hat m}\frac{du}{|\widetilde{\phi_\Delta}(u)|^2}-\frac{\kappa}{2\pi n\tilde q_\Delta}\int_{-\pi\hat m}^{\pi\hat m}\frac{du}{|\widetilde{\phi_\Delta}(u)|^2}\right)\mathbf{1}_\Omega\right]\\ &\quad+\mathbb{E}\left[\left(\frac{96}{2\pi n\tilde q_\Delta}\int_{-\pi\hat m}^{\pi\hat m}\left|\frac{1}{\widetilde{\phi_\Delta}(u)}-\frac{1}{\phi_\Delta(u)}\right|^2du\right)\mathbf{1}_\Omega\right].\end{aligned}$$

Now we choose $\kappa\geq 96$ (which makes the first r.h.s. difference negative or zero), use that on Ω, $1/\tilde q_\Delta\leq(3/2)(1/q_\Delta)$ and that $\hat m\leq n\Delta$ which, together with Lemma 5.1 implies that $\mathbb{E}\left([24\mathrm{pen}(\hat m)-\widehat{\mathrm{pen}}(\hat m)]\mathbf{1}_\Omega\right)\leq c/n$. Plugging this in (30) yields, for $\kappa\geq\kappa_0=96$ that, $\forall m\in\mathcal{M}_n$,

$$\mathbb{E}(\|\hat f_{\hat m}-f\|^2\mathbf{1}_\Omega)\leq 3\|f-f_m\|^2+6(\kappa+8)\mathrm{pen}(m)+\frac{c}{n\Delta}.$$

On the other hand

$$\begin{aligned}\mathbb{P}(\Omega^c) &= \mathbb{P}\left(\left|\frac{1}{\tilde{q}_\Delta}-\frac{1}{q_\Delta}\right|>\frac{1}{2}\frac{1}{q_\Delta}\right)\leq (2q_\Delta)^6\,\mathbb{E}\left(\left|\frac{1}{\tilde{q}_\Delta}-\frac{1}{q_\Delta}\right|^6\right)\\ &\leq (2q_\Delta)^6 C'(3,\Delta)\frac{1}{\Delta^9 n^3}\leq \frac{c}{(n\Delta)^3},\end{aligned}$$

where the last line follows from Proposition 2.1. Moreover $\|\hat{f}_{\hat{m}}-f\|^2=\|\hat{f}_{\hat{m}}-f_{\hat{m}}\|^2+\|f-f_{\hat{m}}\|^2$. Now, $\|f-f_{\hat{m}}\|^2\leq\|f\|^2$ and as

$$\|\hat{f}_{\hat{m}}-f_{\hat{m}}\|^2\leq\|\hat{f}_{n\Delta}-f_{n\Delta}\|^2\leq c(n\Delta)^2,$$

we obtain that

$$\mathbb{E}(\|\hat{f}_{\hat{m}}-f\|^2\mathbf{I}_{\Omega^c})\leq\frac{c}{n\Delta}.$$

This, together with (30) implies the result given in Theorem 2.1. $\Box$

5.5. **Proof of Lemma 5.2.** The result follows from a standard application of the Talagrand inequality to

$$t\mapsto\nu_n(t)=\frac{1}{2\pi n}\sum_{j=1}^{n}\int_{-\pi(m\vee m')}^{\pi(m\vee m')}t^*(-u)\frac{e^{iuY_j(\Delta)}\mathbf{I}_{Y_j(\Delta)\neq 0}-Q_\Delta(u)}{q_\Delta\phi_\Delta(u)}du.$$

First, we have

$$\mathbb{E}\left(\sup_{t\in S_{m'\vee m},\|t\|=1}|\nu_n(t)|^2\right)\leq\frac{1}{2\pi n}\int_{-\pi(m\vee m')}^{\pi(m\vee m')}\frac{1}{q_\Delta|\phi_\Delta^2(u)|}du:=H^2.$$

Note that as $1\leq 1/|\phi_\Delta(u)|\leq 1+2\mu\Delta$,

$$\frac{1}{q_\Delta}\frac{m\vee m'}{n}\leq H^2\leq\frac{(1+2\mu\Delta)^2}{q_\Delta}\frac{m\vee m'}{n}.$$

Second, using

$$|\phi_\Delta^{-1}(u)|\leq 1+2\mu\Delta\text{ and }f^*(u)=\frac{Q_\Delta(u)}{q_\Delta\phi_\Delta(u)},$$

we get

$$\begin{aligned}&\sup_{t\in S_{m'\vee m},\|t\|=1}\mathrm{Var}\left(\frac{1}{2\pi}\int_{-\pi(m\vee m')}^{\pi(m\vee m')}\frac{t^*(-u)e^{iuY_1(\Delta)}\mathbf{I}_{Y_1(\Delta)\neq 0}}{q_\Delta\phi_\Delta(u)}du\right)\\
&\leq\frac{1}{4\pi^2}\sup_{t\in S_{m'\vee m},\|t\|=1}\mathbb{E}\left(\iint\frac{t^*(-u)t^*(w)e^{i(u-w)Y_1(\Delta)}\mathbf{I}_{Y_1(\Delta)\neq 0}}{q_\Delta^2\phi_\Delta(u)\phi_\Delta(-w)}dudw\right)\\
&\leq\frac{1}{2\pi}\left(\iint_{[-\pi(m\vee m'),\pi(m\vee m')]^2}\frac{|Q_\Delta(u-w)|^2}{q_\Delta^2|\phi_\Delta(u)|^2|\phi_\Delta(w)|^2}dudw\right)^{1/2}\\
&=\frac{1}{2\pi}\left(\iint_{[-\pi(m\vee m'),\pi(m\vee m')]^2}\frac{|Q_\Delta(u-w)|^2}{q_\Delta^2|\phi_\Delta(u-w)|^2}\frac{1}{|\phi_\Delta(u)|^2}\frac{|\phi_\Delta(u-w)|^2}{|\phi_\Delta(w)|^2}dudw\right)^{1/2}\\
&\leq(1+2\mu_1\Delta)\|f\|\left(\frac{1}{2\pi}\int_{-\pi(m\vee m')}^{\pi(m\vee m')}\frac{du}{|\phi_\Delta^2(u)|}\right)^{1/2}:=v.\end{aligned}$$

Note that $v\leq c\sqrt{m\vee m'}$.

Lastly

$$\sup_{t\in S_{m'\vee m},\|t\|=1}\sup_{y\in\mathbb{R}}\left|\int\frac{t^*(-u)e^{iuy}\mathbf{I}_{y\neq 0}}{q_\Delta\phi_\Delta(u)}du\right|\leq\sqrt{2\pi}\left(\int_{-\pi(m\vee m')}^{\pi(m\vee m')}\frac{du}{|q_\Delta\phi_\Delta(u)|^2}\right)^{1/2}\leq c\frac{\sqrt{m\vee m'}}{\Delta}:=M$$

Thus,

$$\frac{v}{n}\propto\frac{\sqrt{m\vee m'}}{n},\quad\frac{nH^2}{v}\propto\frac{\sqrt{m\vee m'}}{\Delta},\quad\frac{M^2}{n^2}\propto\frac{m\vee m'}{(n\Delta)^2},\quad\frac{nH}{M}\propto\sqrt{n\Delta},$$

and the Talagrand inequality with $\epsilon^2=\frac{1}{4}$ gives the result. □

5.6. **Proof of Proposition 2.3.** It follows from Lemma 6.3 (see the Appendix) that

$$|\widetilde{g^*_\Delta}(u)-g^*_\Delta(u)|\leq|\widehat{g^*_\Delta}(u)-g^*_\Delta(u)|=|T_1|+|T_2|$$

with

$$T_1=\frac{1}{\tilde{q}_\Delta}(\hat{Q}_\Delta(u)-Q_\Delta(u)),\qquad T_2=Q_\Delta(u)(\frac{1}{\tilde{q}_\Delta}-\frac{1}{q_\Delta}).$$

As $q_\Delta\leq\frac{\mu_1\Delta}{1+\mu_1\Delta}$ and $\frac{1}{\tilde{q}_\Delta}\leq\frac{2(\mu_0\Delta+1)}{\mu_0\Delta}$, using Inequality (24), we get

$$\mathbb{E}\left(|T_1|^{2v}\right)\leq\frac{C(v,\mu_0,\mu_1)}{(n\Delta)^v}.$$

Second, as $|Q_\Delta(u)|\leq q_\Delta$, we obtain by Proposition 2.1

$$\mathbb{E}\left(|T_2|^{2v}\right)\leq\left(\frac{q_\Delta}{\Delta}\right)^{2v}\frac{C(v,\mu_0,\mu_1)}{(n\Delta)^v}.$$

The proof is now complete. □

5.7. **Proof of Corollary 2.1.** The proof follows the lines of Chesneau *et al.* (2012). For $|v|\leq 1$ and $|w|\leq 1$, we have for any $k\geq 1$,

$$|w^k-v^k|=|(w-v)^k+\sum_{j=1}^{k-1}\binom{k}{j}v^j(w-v)^{k-j}|\leq|w-v|^k+|v|\sum_{j=1}^{k-1}\binom{k}{j}|w-v|^{k-j}$$

Thus,

$$|w^k-v^k|^2\leq 2(|w-v|^{2k}+D_k^2|v|^2|w-v|^2).$$

We apply this inequality for $w=\widetilde{g^*_{m,\Delta}}(u)$, $v=g^*_{m,\Delta}(u)$ and use Proposition 2.3 to obtain:

$$\mathbb{E}|(\widetilde{g^*_{m,\Delta}}(u))^k-(g^*_{m,\Delta}(u))^k|^2\quad\leq\quad 2\left(\frac{C(k,\mu_0,\mu_1)}{(n\Delta)^k}+D_k^2|g^*_\Delta(u)|^2\frac{C(1,\mu_0,\mu_1)}{n\Delta}\right).$$

Finally, the Plancherel theorem and Inequality (16) give

$$\begin{aligned}\mathbb{E}\left(\|\widehat{g^{\star k}_{m,\Delta}}-g^{\star k}_{m,\Delta}\|^2\right)&=\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\mathbb{E}\left(|(\widetilde{g^*_{m,\Delta}}(u))^k-(g^*_{m,\Delta}(u))^k|^2\right)du\\&\leq c(k,\mu_0,\mu_1)\left(\frac{m}{(n\Delta)^k}+D_k^2\frac{\|f\|^2}{n\Delta}\right).\end{aligned}$$

This ends the proof of Corollary 2.1. □

5.8. **Proof of Proposition 2.4.** We have

$$F_k(\tilde{\mu}\Delta) - F_k(\mu\Delta) \quad = \quad (\tilde{\mu} - \mu)\Delta\Big(\sum_{j=0}^{k-1}(\tilde{\mu}\Delta)^j(\mu\Delta)^{k-1-j} + \sum_{j=0}^{k}(\tilde{\mu}\Delta)^j(\mu\Delta)^{k-j}\Big)$$

Then, we use inequality (10), which also implies $\mathbb{E}(|\tilde{\mu}|^{2p}) \leq C(p, \mu_1)$ to get the result. □

5.9. **Proof of Proposition 2.5.** We write

$$\hat{f}_{m,K} = \hat{f}_m^{(1)} + \hat{f}_{m,K}^{(2)}, \quad \hat{f}_m^{(1)} = (1 + \tilde{\mu}\Delta)\widehat{g_{m,\Delta}}.$$

We define

$$\tilde{f}_{m,K}^{(2)} = \sum_{k=1}^{K}(-1)^k F_k(\tilde{\mu}\Delta) g_{m,\Delta}^{\star\ k+1}, \quad f_{m,K} = \sum_{k=0}^{K}(-1)^k F_k(\mu\Delta) g_{m,\Delta}^{\star\ k+1},$$

and

$$f_m^{(1)} = (1 + \mu\Delta) g_{m,\Delta}, \quad f_{m,K}^{(2)} = f_{m,K} - f_m^{(1)}.$$

By the triangle inequality we have

$$\begin{aligned}
\mathbb{E}(\|\hat{f}_{m,K} - f\|^2) \quad \leq \quad & 4\Big(\mathbb{E}(\|\hat{f}_m^{(1)} - f_m^{(1)}\|^2) + \mathbb{E}(\|\hat{f}_{m,K}^{(2)} - \tilde{f}_{m,K}^{(2)}\|^2) + \mathbb{E}(\|\tilde{f}_{m,K}^{(2)} - f_{m,K}^{(2)}\|^2) \\
& + \|f_{m,K} - f_m\|^2\Big) + \|f_m - f\|^2 \\
:= \ & 4(T_1 + T_2 + T_3 + T_4) + \|f_m - f\|^2,
\end{aligned}$$

where $\|f_m - f\|^2$ is the usual bias term

$$\|f_m - f\|^2 = \frac{1}{2\pi}\int_{|u|\geq \pi m} |f^*(u)|^2 du,$$

Now we successively study the terms T_i, for $i = 1, \ldots, 4$.

Let us start by the study of T_1. We split it again $T_1 \leq 3(T_{1,1} + T_{1,2} + T_{1,3})$ with

$$T_{1,1} = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{1 + \tilde{\mu}\Delta}{\tilde{q}_\Delta} - \frac{1 + \mu\Delta}{q_\Delta}\right|^2 |\hat{Q}_\Delta(u) - Q_\Delta(u)|^2 du$$

$$T_{1,2} = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{1 + \tilde{\mu}\Delta}{\tilde{q}_\Delta} - \frac{1 + \mu\Delta}{q_\Delta}\right|^2 |Q_\Delta(u)|^2 du, T_{1,3} = \frac{1}{2\pi}\left(\frac{1 + \mu\Delta}{q_\Delta}\right)^2\int_{-\pi m}^{\pi m}|\hat{Q}_\Delta(u) - Q_\Delta(u)|^2 du.$$

As $\mathbb{E}(|\hat{Q}_\Delta(u) - Q_\Delta(u)|^2) \leq q_\Delta/n$, we find

$$\mathbb{E}(T_{1,3}) \leq \frac{(1 + \mu\Delta)^2}{q_\Delta}\frac{m}{n} = \frac{(1 + \mu\Delta)^3}{\mu}\frac{m}{n\Delta}.$$

This term is the main one. Now we prove that the two others are of order less than $O(1/(n\Delta))$.

First, it follows from Proposition 2.1 that

$$\mathbb{E}\left(\left|\frac{1 + \tilde{\mu}\Delta}{\tilde{q}_\Delta} - \frac{1 + \mu\Delta}{q_\Delta}\right|^{2k}\right) \leq C(k, \mu_0, \mu_1)\frac{1}{\Delta^{2k}}\frac{1}{(n\Delta)^k}. \tag{31}$$

Moreover, using (24), we have

$$\mathbb{E}\left(|\hat{Q}_\Delta(u) - Q_\Delta(u)|^4\right) \leq c(2, \mu_1)\left(\frac{\Delta}{n}\right)^2.$$

Therefore, by the Cauchy-Schwarz Inequality and assuming that $n\Delta \geq 1$ and $m \leq n\Delta$, we obtain that

$$\mathbb{E}(T_{1,1}) \leq c(\mu_0, \mu_1)\frac{m}{(n\Delta)^2} \leq \frac{c(\mu_0, \mu_1)}{n\Delta}.$$

Now, using (31) with $k=1$, and that $|Q_\Delta(u)| \le q_\Delta |f^*(u)|$, we get

$$\mathbb{E}(T_{1,2}) \le c_1(\mu_0,\mu_1)\frac{(q_\Delta/\Delta)^2}{n\Delta}\frac{1}{2\pi}\int_{-\pi m}^{\pi m}|f^*(u)|^2 du \le \frac{c(\mu_0,\mu_1)\|f\|^2}{n\Delta}.$$

Gathering the terms, we have

$$\mathbb{E}(T_1) \le 3\frac{(1+\mu\Delta)^3}{\mu}\frac{m}{n\Delta} + \frac{c(\mu_0,\mu_1,\|f\|)}{n\Delta}.$$

Next we study T_2:

$$\begin{aligned}\mathbb{E}\big(\|\hat{f}^{(2)}_{m,K} - \tilde{f}^{(2)}_{m,K}\|^2\big) &= \mathbb{E}\left(\Big\|\sum_{k=1}^{K}(-1)^k(F_k(\mu\Delta) - F_k(\tilde{\mu}\Delta) + F_k(\mu\Delta))\big(\widehat{g^{\star\ k+1}_{m,\Delta}} - g^{\star\ k+1}_{m,\Delta}\big)\Big\|^2\right)\\ &\le c(K,\mu_1)\sum_{k=1}^{K}\Delta^{2k}\mathbb{E}\left(\|\widehat{g^{\star k+1}_{m,\Delta}} - g^{\star k+1}_{m,\Delta}\|^2\right).\end{aligned}$$

From Proposition 2.1, as $m \le n\Delta$, this term is bounded by

$$\sum_{k=1}^{K}\Delta^{2k}c(k+1,\mu_0,\mu_1)\left(\frac{m}{(n\Delta)^{k+1}} + \frac{D^2_{k+1}}{n\Delta}\|f\|^2\right) \le \frac{c(K,\mu_0,\mu_1)}{n\Delta}.$$

Now, to bound T_3, we write that, using that $|g^*_\Delta(u)| \le 1 \wedge |f^*(u)|$, for all u,

$$\begin{aligned}2\pi\|\tilde{f}^{(2)}_{m,K} - f^{(2)}_{m,K}\|^2 &\le \int_{-\pi m}^{\pi m}\Big|\sum_{k=1}^{K}(-1)^k\big(F_k(\tilde{\mu}\Delta) - F_k(\mu\Delta)\big)\big(g^*_\Delta(u)\big)^{k+1}\Big|^2 du\\ &\le \|f^*\|^2\Big|\sum_{k=1}^{K}\big|F_k(\tilde{\mu}\Delta) - F_k(\mu\Delta)\big|\Big|^2.\end{aligned}$$

Then, with Proposition 2.4 we get

$$\mathbb{E}\left(\|\tilde{f}^{(2)}_{m,K} - f^{(2)}_{m,K}\|^2\right) \le K\|f\|^2\sum_{k=1}^{K}\mathbb{E}\left(|F_k(\tilde{\mu}\Delta) - F_k(\mu\Delta)|^2\right) \le c(K,\mu_1)\frac{\|f\|^2}{n\Delta}.$$

Thus T_3 is bounded by a term of order $1/(n\Delta)$.
Next we turn to T_4. We have, using that $|g^*_\Delta(u)| \le 1 \wedge |f^*(u)|$ for all u,

$$\begin{aligned}2\pi\|f_{m,K} - f_m\|^2 = \|f^*_{m,K} - f^*_m\|^2 &\le \int_{-\pi m}^{\pi m}\Big|\sum_{k=K+1}^{\infty}(1+\mu_1\Delta)(\mu_1\Delta)^{k+1}\big(g^*_\Delta(u)\big)^{k+1}\Big|^2 du\\ &\le \|f^*\|^2\Big(\sum_{k=K+1}^{\infty}(1+\mu_1\Delta)(\mu_1\Delta)^{k+1}\Big)^2\\ &= \frac{(1+\mu_1\Delta)^2}{(1-\mu_1\Delta)^2}(\mu_1\Delta)^{2K+2}\|f^*\|^2 \le 2\pi A(\mu_1\Delta)^{2K+2},\end{aligned}$$

where A is defined in Proposition 2.1. It follows that,

$$\mathbb{E}(\|\hat{f}_{m,K} - f\|^2) \le \|f_m - f\|^2 + 4\Big(A(\mu_1\Delta)^{2K+2} + 3\frac{(1+\mu\Delta)^3}{\mu}\frac{m}{n\Delta} + E_K\frac{1}{n\Delta}\Big),$$

which is the result of Proposition 2.1. $\square$

5.10. **Proof of Theorem 2.2.** The only term involved in the adaptation is $T_{1,3}$, all the others are bounded independently of m. Thus the above bounds can be used, and $T_{1,3}$ is treated by using Talagrand Inequality applied to the underlying empirical process. The principle is as in Comte *et al.* (2014), proof of Theorem 4.1, with the use of the set Ω as in the proof of Theorem 2.1. For sake of conciseness, the proof is omitted.

5.11. **Proofs of Section 3.** We start by stating a useful Lemma.

Lemma 5.3. *Assume that* **[B]** (i)-(ii)(p) *hold. Then*

$$\mathbb{E}\left(\left|\widetilde{\psi}(v)-\psi(v)\right|^{2p}\right)\leq\frac{\kappa_p}{(n\Delta)^p}\frac{1+|\psi(v)|^{2p}}{|H_\Delta(v)|^{2p}}.$$

Proof of Lemma 5.3.
We omit the index Δ for simplicity. We have the decomposition

$$\widetilde{\psi}-\psi=\frac{\hat G}{\tilde H}-\frac{G}{H}=(\hat G-G)\left(\frac{1}{\tilde H}-\frac{1}{H}\right)+\frac{\hat G-G}{H}+G\left(\frac{1}{\tilde H}-\frac{1}{H}\right)\tag{32}$$

so that a bound follows from bounding $\mathbb{E}(|\hat G(v)-G(v)|^2)$ and $\mathbb{E}\left(|\tilde H^{-1}(v)-H^{-1}(v)|^2\right)$.
Clearly

$$\mathbb{E}(|\hat G(v)-G(v)|^2)=\frac{1}{n\Delta^2}\mathrm{Var}(Y_1(\Delta)e^{ivY_1(\Delta)})\leq\frac{\mathbb{E}(|Y_1(\Delta)|^2/\Delta)}{n\Delta},\tag{33}$$

where $\mathbb{E}(|Y_1(\Delta)|^2/\Delta)=\mathbb{E}\Lambda\mathbb{E}\xi^2+\Delta\mathbb{E}(\Lambda^2)(\mathbb{E}(\xi))^2$. And for general p, the Rosenthal Inequality yields

$$\begin{aligned}\mathbb{E}(|\hat G(v)-G(v)|^{2p}) &\leq \frac{C(2p)}{(n\Delta)^{2p}}\left\{n2^{2p}\mathbb{E}(|Y_1(\Delta)|^{2p})+[n\mathrm{Var}(Y_1(\Delta)e^{ivY_1(\Delta)})]^p\right\}\\ &\leq c(\frac{1}{(n\Delta)^{2p-1}}+\frac{1}{(n\Delta)^p})\end{aligned}\tag{34}$$

since **[B]**(ii)(p) holds and $\Delta\leq 1$.

Next the bound on $\mathbb{E}\left(|\tilde H^{-1}-H^{-1}|^{2p}\right)$ is the following:

$$\mathbb{E}\left(\left|\frac{1}{\tilde H(v)}-\frac{1}{H(v)}\right|^{2p}\right)\leq c_p\inf\left((n\Delta)^{-p}|H(v)|^{-4p},|H(v)|^{-2p}\right).\tag{35}$$

The proof of (35) is similar to the proof of Lemma 5.1 and thus is omitted. We conclude using (32), (33), (34) and (35). □

Proof of Proposition 3.1.
We first prove inequality (20) using Lemma 6.3 (see the Appendix). Let

$$R(u)=\int_0^u\left(\frac{\hat G(v)}{\tilde H(v)}-\frac{G(v)}{H(v)}\right)dv.$$

Then to compute the risk of the estimator, we write:

$$\|\widetilde{f}_m-f\|^2=\|f-f_m\|^2+\|\widetilde{f}_m-f_m\|^2=\|f-f_m\|^2+\frac{1}{2\pi}\int_{-\pi m}^{\pi m}|\widetilde{f^*}(u)-f^*(u)|^2du$$

and

$$\begin{aligned}|\widetilde{f^*}(u)-f^*(u)|^2 &\leq |\widetilde{f^*}(u)-f^*(u)|^2\mathbf{1}_{|R(u)|<1}+|\widetilde{f^*}(u)-f^*(u)|^2\mathbf{1}_{|R(u)|\geq 1}\\ &\leq |\widehat{f^*}(u)-f^*(u)|^2\mathbf{1}_{|R(u)|<1}+4\mathbf{1}_{|R(u)|\geq 1}\\ &\leq |f^*(u)|^2\left|\exp(R(u))-1\right|^2\mathbf{1}_{|R(u)|<1}+4\mathbf{1}_{|R(u)|\geq 1}\\ &\leq e^2|f^*(u)|^2|R(u)|^2+4|R(u)|^{2p}\end{aligned}$$

By the Hölder Inequality, $\mathbb{E}(|R(u)|^{2p})\leq |u|^{2p-1}\int_0^u \mathbb{E}(|\widetilde{\psi}(v)-\psi(v)|^{2p})dv$. Then Lemma 5.3 gives the result (20).

Now we prove inequality (21). We have

$$|\widetilde{f^*}(u)-f^*(u)|^2\leq e^2|f^*(u)|^2|R(u)|^2\mathbf{1}_{|R(u)|\leq 1}+4|R(u)|\mathbf{1}_{|R(u)|>1}.$$

So we prove

$$\mathbb{E}(|R(u)|^2\mathbf{1}_{|R(u)|\leq 1})\leq \frac{c}{n\Delta}\left(M_1\int_0^{|u|}\frac{dv}{|H(v)|^2}+\left(\int_0^{|u|}\frac{|G(v)|}{|H(v)|^2}dv\right)^2\right) \tag{36}$$

and

$$\begin{aligned}\mathbb{E}(|R(u)|\mathbf{1}_{|R(u)|>1}) \leq\ & c\left\{\left(M_p+\left(\int_0^{|u|}\left|\frac{G(v)}{H(v)}\right|^2dv\right)^p\right)\left(\frac{1}{n\Delta}\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^p\right.\\ &\left.+\frac{\mathbb{E}(|Y_1(\Delta)|^{2p}/\Delta)}{(n\Delta)^{2p-1}}\left(\int_0^{|u|}\frac{dv}{|H(v)|}\right)^{2p}\right\}\end{aligned} \tag{37}$$

with $M_p=[\|G'\|_1^p+\mathbb{E}^{1/2}(Y_1^{2p}(\Delta)/\Delta)]$.

By decomposition (32), we write that $|R(u)|\leq R_1(u)+R_2(u)+R_3(u)$ with

$$R_1(u)=\left|\int_0^u\frac{\hat{G}(v)-G(v)}{H(v)}dv\right|,\quad R_2(u)=\left|\int_0^u G(v)\left(\frac{1}{\tilde{H}(v)}-\frac{1}{H(v)}\right)dv\right|$$

and

$$R_3(u)=\left|\int_0^u(\hat{G}(v)-G(v))\left(\frac{1}{\tilde{H}(v)}-\frac{1}{H(v)}\right)dv\right|.$$

Let $A_j:=\{|R(u)|\leq 1\}\cap\{\arg\max_{k\in\{1,2,3\}}R_k(u)=j\}$, then

$$\begin{aligned}\mathbb{E}(|R(u)|^2\mathbf{1}_{|R(u)|\leq 1}) &\leq 9\mathbb{E}(R_1^2(u)\mathbf{1}_{A_1})+9\mathbb{E}(R_2^2(u)\mathbf{1}_{A_2})+\mathbb{E}(|R(u)|\mathbf{1}_{A_3})\\ &\leq 9\mathbb{E}(R_1^2(u))+9\mathbb{E}(R_2^2(u))+3\mathbb{E}(R_3(u)\mathbf{1}_{A_3}).\end{aligned} \tag{38}$$

Then

$$\begin{aligned}\mathbb{E}(R_1^2(u)) &\leq \frac{1}{n\Delta^2}\int_0^u\int_0^u\frac{\mathbb{E}(Y_1^2(\Delta)e^{i(v-w)Y_1(\Delta)})}{H(v)H(-w)}dvdw\\ &\leq \frac{1}{n\Delta}\left(\int_0^u\int_0^u\frac{1}{|H(v)|^2}|G'(v-w)|dvdw\right)^{1/2}\left(\int_0^u\int_0^u\frac{1}{|H(w)|^2}|G'(v-w)|dvdw\right)^{1/2}\\ &\leq \frac{\|G'\|_1}{n\Delta}\int_0^u\frac{dv}{|H(v)|^2}.\end{aligned} \tag{39}$$

Moreover

$$\begin{aligned}
\mathbb{E}(R_2^2(u)) &\leq \int_0^u \int_0^u G(v)G(-w)\mathbb{E}\left[\left(\frac{1}{\tilde{H}(v)} - \frac{1}{H(v)}\right)\left(\frac{1}{\tilde{H}(-w)} - \frac{1}{H(-w)}\right)\right] dvdw \\
&\leq \frac{c}{n\Delta}\int_0^u \int_0^u \frac{|G(v)G(-w)|}{|H(v)|^2|H(w)|^2} dvdw = \frac{c}{n\Delta}\left(\int_0^u \frac{|G(v)|}{|H(v)|^2} dv\right)^2 \qquad (40)
\end{aligned}$$

Lastly

$$\begin{aligned}
\mathbb{E}(R_3(u)) &\leq \int_0^u \mathbb{E}^{1/2}(|\hat{G}(v) - G(v)|^2)\mathbb{E}^{1/2}\left(\left|\frac{1}{\tilde{H}(v)} - \frac{1}{H(v)}\right|^2\right) dv \\
&\leq c\frac{\mathbb{E}^{1/2}(Y_1^2(\Delta)/\Delta)}{n\Delta}\int_0^u \frac{1}{|H(v)|^2} dv \qquad (41)
\end{aligned}$$

We plug (39)-(41) in (38) and we obtain (36).

Let now $B_j := \{|R(u)| > 1\} \cap \{\arg\max_{k\in\{1,2,3\}} R_k(u) = j\}$. On B_j, $|R(u)| \leq 3R_j(u)$ and thus $3R_j(u) > 1$. Then

$$\begin{aligned}
\mathbb{E}(|R(u)|\mathbf{1}_{|R(u)|>1}) &\leq 3(\mathbb{E}(R_1(u)\mathbf{1}_{B_1}) + \mathbb{E}(R_2(u)\mathbf{1}_{B_2}) + \mathbb{E}(R_3(u)\mathbf{1}_{B_3})) \\
&\leq 9^p(\mathbb{E}(R_1^{2p}(u)) + \mathbb{E}(R_2^{2p}(u))) + 3^p\mathbb{E}(R_3^p(u)\mathbf{1}_{B_3}). \qquad (42)
\end{aligned}$$

By applying Rosenthal's inequality and using the bound obtained in (39), we get

$$\mathbb{E}(R_1^{2p}(u)) \leq c\left(\|G'\|_1^p\left(\frac{1}{n\Delta}\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^p + \frac{\mathbb{E}(|Y_1(\Delta)|^{2p}/\Delta)}{(n\Delta)^{2p-1}}\left(\int_0^{|u|}\frac{dv}{|H(v)|}\right)^{2p}\right).$$

For R_2 we write

$$\mathbb{E}(R_2^{2p}(u)) \leq \left(\int_0^{|u|}\left(\frac{|G(v)|}{|H(v)|}\right)^2 dv\right)^p \mathbb{E}\left[\left(\int_0^{|u|}|H(v)|^2\left|\frac{1}{\tilde{H}(v)} - \frac{1}{H(v)}\right|^2\right)^p\right].$$

Now we apply the Hölder inequality and inequality (35),

$$\begin{aligned}
\mathbb{E}(R_2^{2p}(u)) &\leq \left(\int_0^{|u|}\left(\frac{|G(v)|}{|H(v)|}\right)^2 dv\right)^p\left(\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^{p-1}\int_0^{|u|}\frac{|H(v)|^{4p}}{|H(v)|^2}\mathbb{E}\left(\left|\frac{1}{\tilde{H}(v)} - \frac{1}{H(v)}\right|^{2p}\right) dv \\
&\leq c\left(\int_0^{|u|}\left(\frac{|G(v)|}{|H(v)|}\right)^2 dv\right)^p\left(\frac{1}{n\Delta}\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^p.
\end{aligned}$$

For R_3 we apply the Hölder Inequality again, and then the Cauchy Schwarz Inequality, (34) and (35), to obtain

$$\begin{aligned}
\mathbb{E}(R_3^p(u)) &\leq \left(\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^{p-1}\int_0^{|u|}\frac{|H(v)|^{2p}}{|H(v)|^2}\mathbb{E}\left(\left(|\hat{G}(v) - G(v)|\left|\frac{1}{\tilde{H}(v)} - \frac{1}{H(v)}\right|\right)^p\right) dv \\
&\leq \left(\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^{p-1}\int_0^{|u|}\frac{|H(v)|^{2p}}{|H(v)|^2}\mathbb{E}^{1/2}\left(|\hat{G}(v) - G(v)|^{2p}\right)\mathbb{E}^{1/2}\left(\left|\frac{1}{\tilde{H}(v)} - \frac{1}{H(v)}\right|^{2p}\right) dv \\
&\leq c\,\mathbb{E}^{1/2}(|Y_1(\Delta)|^{2p}/\Delta)\left(\frac{1}{n\Delta}\int_0^{|u|}\frac{dv}{|H(v)|^2}\right)^p.
\end{aligned}$$

Plugging the three bounds in (42) gives (37). □

6. Appendix

The Talagrand inequality. The result below follows from the Talagrand concentration inequality given in Klein and Rio (2005) and arguments in Birgé and Massart (1998) (see the proof of their Corollary 2 page 354).

Lemma 6.1. *(Talagrand Inequality) Let* $Y_1, \dots, Y_n$ *be independent random variables, let* $\nu_{n,Y}(f) = (1/n)\sum_{i=1}^n [f(Y_i) - \mathbb{E}(f(Y_i))]$ *and let* $\mathcal{F}$ *be a countable class of uniformly bounded measurable functions. Then for* $\epsilon^2 > 0$

$$\mathbb{E}\Big[\sup_{f\in\mathcal{F}} |\nu_{n,Y}(f)|^2 - 2(1+2\epsilon^2)H^2\Big]_+ \leq \frac{4}{K_1}\left(\frac{v}{n}e^{-K_1\epsilon^2\frac{nH^2}{v}} + \frac{98M^2}{K_1 n^2 C^2(\epsilon^2)}e^{-\frac{2K_1C(\epsilon^2)\epsilon}{7\sqrt{2}}\frac{nH}{M}}\right),$$

with $C(\epsilon^2) = \sqrt{1+\epsilon^2} - 1$, $K_1 = 1/6$, *and*

$$\sup_{f\in\mathcal{F}} \|f\|_\infty \leq M, \quad \mathbb{E}\Big[\sup_{f\in\mathcal{F}} |\nu_{n,Y}(f)|\Big] \leq H, \ \sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{k=1}^n \mathrm{Var}(f(Y_k)) \leq v.$$

By standard density arguments, this result can be extended to the case where $\mathcal{F}$ is a unit ball of a linear normed space, after checking that $f \mapsto \nu_n(f)$ is continuous and $\mathcal{F}$ contains a countable dense family.

The Rosenthal inequality. (see *e.g.* Hall and Heyde (1980, p.23)) Let $(X_i)_{1\leq i\leq n}$ be n independent centered random variables, such that $\mathbb{E}(|X_i|^p) < +\infty$ for an integer $p \geq 1$. Then there exists a constant $C(p)$ such that

$$\mathbb{E}\left(\left|\sum_{i=1}^n X_i\right|^p\right) \leq C(p)\left(\sum_{i=1}^n \mathbb{E}(|X_i|^p) + \left(\sum_{i=1}^n \mathbb{E}(X_i^2)\right)^{p/2}\right). \tag{43}$$

Lemma 6.2. *Consider* c, s *nonnegative real numbers, and* γ *a real such that* $2\gamma > -1$ *if* $c = 0$ *or* $s = 0$. *Then, for all* $m > 0$,

- $\int_{-m}^m (x^2+1)^\gamma \exp(c|x|^s)dx \approx m^{2\gamma+1-s}e^{cm^s}$,

and if in addition $2\gamma > 1$ *if* $c = 0$ *or* $s = 0$,

- $\int_m^\infty (x^2+1)^{-\gamma} \exp(-c|x|^s)dx \approx m^{-2\gamma+1-s}e^{-cm^s}$.

The proof of this lemma is based on integration by parts and is omitted. See also Lemma 2 p. 35 in Butucea and Tsybakov (2008a,b).

Finally we state a useful and elementary inequality.

Lemma 6.3. *Let* $z = re^{i\theta}$ *with* $r \leq 1$, *and* $\hat{z} = \rho e^{i\omega}, \tilde{z} = e^{i\omega}$ *with* $\rho > 1$. *Then* $|\tilde{z} - z| \leq |\hat{z} - z|$.

References

[1] Antonio, K. and Beirlant, J. (2007) Actuarial statistics with generalized linear mixed models. *Insurance Math. Econom.* **40**, 58-76.

[2] Belomestny, D. (2011) Statistical inference for time-changed Lévy processes via composite characteristic function estimation. *Ann. Statist.* **39**, 2205-2242.

[3] Belomestny, D. and Schoenmakers, J (2014) Statistical Skohorod embedding problem and its generalizations. Working paper.

[4] Birgé, L. and Massart, P. (1998) Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* **4**, 329-375.

[5] Birgé, L. and Massart, P. (2007) Minimal penalties for Gaussian model selection. *Probab. Theory Related Fields* **138**, 33-73.

[6] Butucea, C. and Tsybakov, A. B. (2008a) Sharp optimality in density deconvolution with dominating bias. I. *Theory Probab. Appl.* **52**, 24-39.

[7] Butucea, C. and Tsybakov, A. B. (2008b) Sharp optimality in density deconvolution with dominating bias. II. *Theory Probab. Appl.* **52**, 237-249.

[8] Chesneau, C., Comte, F. and Navarro, F. (2013) Fast nonparametric estimation for convolutions of densities. *the Canadian Journal of Statistics*, **41**, 617-636.

[9] Comte, F., Duval, C. and Genon-Catalot, V. (2014) Nonparametric density estimation in compound Poisson processes using convolution power estimators. *Metrika* **77**, 163-183.

[10] Comte, F. and Genon-Catalot (2013) Adaptive Laguerre density estimation for mixed Poisson models. Preprint Hal, Prépublication MAP5 2013-15.

[11] Comte, F. and Genon-Catalot (2009) Nonparametric estimation for pure jump Lévy processes based on high frequency data. *Stochastic Processes and their Applications* **119**, 4088-4123.

[12] Davidian, M., and Giltinan, D.M. (1995) *Nonlinear Models for Repeated Measurement Data*, New York: Chapman and Hall.

[13] Duval, C. (2013) Density estimation for compound Poisson processes from discrete data. *Stochastic Process. Appl.* **123**, 3963-3986.

[14] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling extremal events.* Berlin: Springer.

[15] Figueroa-López, J. E. (2009) Nonparametric estimation of Lévy models based on discrete-sampling. Optimality, 117-146, IMS Lecture Notes Monogr. Ser., 57, Inst. Math. Statist., Beachwood, OH.

[16] van Es, B., Gugushvili, S. and Spreij, P. (2007) A kernel type nonparametric density estimator for decompounding. *Bernoulli* **13**, 672694.

[17] Grandell, J. (1997) *Mixed Poisson processes.* Monographs on Statistics and Applied Probability, 77. Chapman & Hall, London.

[18] Gugushvili, S. (2012) Nonparametric inference for discretely sampled Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **48**, 282-307.

[19] Hall, P. and Heyde, C. C. (1980) *Martingale limit theory and its application.* Probability and Mathematical Statistics. Academic Press, New York-London.

[20] Klein, T. and Rio, E. (2005) Concentration around the mean for maxima of empirical processes. *Ann. Probab.* **33** 1060-1077.

[21] Mikosch, T. (2009) *Non-life insurance mathematics. An introduction with the Poisson process.* Second edition. Universitext. Springer-Verlag, Berlin.

[22] Neumann, Michael H. and Reiss, M. (2009) Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli* **15**, 223-248.

[23] Pinheiro, J.C. and Bates, D.M. (2000) *Mixed-Effects Models in S and S-PLUS*. New York, Springer.

[24] Ueltzhöfer, F. A. J. and Klüppelberg, C. (2011) An oracle inequality for penalised projection estimation of Lévy densities from high-frequency observations. *J. Nonparametr. Stat.* **23**, 967-989.